Acta Univ. M. Belii
Math. no. 8(2000), pp. 3-9

TANGENTIAL PROLONGATION OF SURFACES IN E_3- CLASSIFICATION OF PARAMETRIC NETS

Anton Dekrét and Ján Bakša

Abstract. The paper is devoted to some applications of the classical differential geometry of surfaces in E_3 in the computer grafics. It contains a classification of parametric nets on surfaces by a tangential prolongation of nets. This classification gives possibilities of the choice of a suitable parametrisation from the point of view drawing of surfaces in computer grafics.

Introduction

In the computer grafics of surfaces in E_3 a very useful frame tool is a suitable parametric net. Let us introduce basic notions on parametric representations of surfaces in E_3, see for example [1], [2]. Let

$$\bar{r}(u,v) = (x(u,v), y(u,v), z(u,v)), (u,v) \in \Omega \subset R^2 \tag{1}$$

be an equation of a surface $\mathcal{P}$ in a cartesian coordinate frame. Let functions $x(u,v)$, $y(u,v)$, $z(u,v)$ be differentiable up to second order. Both the point $(u,v) \in \Omega$ and its image $\bar{r}(u,v)$ on the surface $\mathcal{P}$ will be called regular if the vector product $\bar{r}_u \times \bar{r}_v$ is not equal to zero, where we use the shortened notations

$$\bar{r}_u = \frac{\partial \bar{r}}{\partial u}, \ \bar{r}_v = \frac{\partial \bar{r}}{\partial v}, \ \bar{r}_{uv} = \frac{\partial^2 \bar{r}}{\partial u \partial v}.$$

In the opposite case we will say that points are singular. We suppose that the surface has only a finite number of the singular points.

The curve $\bar{r}(u, v_0)$ or $\bar{r}(u_0, v)$ on the surface $\mathcal{P}$ will be called the parametric $u-$curve going through a point $\bar{r}(u_0, v_0)$ on $\mathcal{P}$. We are interested in surfaces of lines wich are determined by the tangent vectors of the certain parametric curves (for example of the $v-$curves) at the points of a parametric curve of the other type (an $u-$curve). Their equations are as follows

(2) $\bar{R}(u,t,v_0) = \bar{r}(u,v_0) + t\bar{r}_v(u,v_0)$ *or*
(3) $\bar{R}(v,t,u_0) = \bar{r}(u_0,v) + t\bar{r}_u(u_0,v)$.

2000 *Mathematics Subject Classification.* 53C05.

Key words and phrases. Surface in E_3, parametric nets on surfaces, tangential prolongations of nets.

Supporded by the VEGA SR No.1/5011/98.

The surface (2) will be called the tangential $v=$ surface along the $u-$curve for $v=v_0$. Analogously we will say that the surface (3) is the tangential $u-$surface along the $v-$curve for $u=u_0$. We introduce some applications of the tangential $v-$surfaces:

a) $\bar{R}\,(u,t,v_0)$, $t\in< -a,a >\subset R$ is the so-called tangential $v-$belt. It can be used as a grafic information about the surface $\mathcal{P}$ along the $u-$curve for $v=v_0$.

b) The map $(u,v)\mapsto \bar{r}\,(u,v)$ determines a two-parametric mouvment in E_3. Then the mouvments given by the parametric curves can be called "basic mouvments". Then $\bar{R}(u,t,v_0)$, $t\in< 0,1 >$, is the surface of $v-$velocities along the $u-$curve $v=v_0$.

c) $\bar{R}\,(u,t,v_0)$, $t\in< 0,a >$ or $t\in< -a,0 >$, will be called the tangential $v-$prolongation of the surface $\mathcal{P}$ along the $u-$curve for $v=v_0$. Its technical application is clear from the definition.

Analogously it can be said in the case of $u-$surfaces.

The main goal of the paper is the classification of the parametric nets on surfaces in E_3 based on the tangential $v-$ or $u-$ surfaces.

Classification of the parametric nets on surfaces in E_3

The tangential $v-$ or $u-$ surfaces which we have introduced are line surfaces. These ones can by classify as developable (surfaces of tangents of curves, cone - surface of one-parameter family of lines with a steady vertex, cylindre-surface of one-parameter family of parallel lines) and as undevelopable.

Proposition 1. *The tangential $v-$ surface along an $u-$curve is developable if and only if $\bar{r}_{uv}=\alpha\bar{r}_{\hat{u}}+\beta\bar{r}_v$ at every point of the u-curve.*

Proof. Arbitrary line surface $\bar{r}=\bar{a}(u)+t\bar{b}(u)$ is developable if and only if the tangential planes at points of any line are identified, i.e if and only if $\bar{b}_u\cdot(\bar{b}\times\bar{a}_u)=0$. Considering our $v-$ surface in the form (2), i.e. $\bar{a}(u)=\bar{r}(u,v_0)$, $\bar{b}=\bar{r}_v(u,v_0)$, we inmediately get that $\bar{r}_{uv}=\alpha\bar{r}_u+\beta\bar{r}_v$ is the necessary and sufficient condition for the surface $\bar{R}(u,t,v_0)$ to be developable.

Let (2) be the developable tangential $v-$surface, $\bar{r}_{uv}=\alpha\bar{r}_u+\beta\bar{r}_v$. Then its curve

$$\bar{R}^*(u)=\bar{r}\,(u,v_0)+g(u)\bar{r}_v\,(u,v_0) \tag{4}$$

is its edge of regression, i.e. the surface $\bar{R}\,(u,t,v_0)$ is the surface of the tangents of this curve, if and only if $\bar{R}^*_u=k\bar{r}_v\,(u,v_0), i.e.iff$

$$1+g\alpha=0, g_u+g\beta=k. \tag{5}$$

It immediately gives 1.) If $\alpha=0$ then the tangent $v-$surface is a cylinder surface.

2.) If $\alpha \neq 0$, (then $g = -1/\alpha$), and $k = 0, (i.e. \alpha_u = \beta\alpha)$, then the tangent $v-$surface is a cone surface.

3.) If $\alpha \neq 0, (g = -1/\alpha)$ and $\alpha_u \neq \beta\alpha$ then the tangent $v-$surface is the surface of all tangents of the curve (4). The analogous assertions are right for the tangent $u-$surfaces.

We obtain the following classification of parametric nets of surfaces in E_3:

I. Tangent $v-$ and $u-$ surfaces are developable, i.e. $\bar{r}_{uv} = \alpha\bar{r}_u + \beta\bar{r}_v$.

a) $\alpha = 0,\ \beta = 0, i.e., \bar{r}_{uv} = \bar{0}$. Then $\bar{r}(u,v) = \bar{r}_1(u) + \bar{r}_2(v), i.e.$ the surface $\mathcal{P}$ can be created by the translation of a parametric curve along the other one. Both tangent $v-$and $u-$ surface are cylinder surfaces.

b) $\alpha \neq 0,\ \beta = 0$. The tangential $u-$surfaces along $v-$curves are cylinder surfaces. If $\alpha_u = 0$ then the tangent $v-$surfaces along $u-$curves are cones. If $\alpha_u \neq 0$ then the tangent $v-$surfaces along $u-$curves are surfaces of tangents of curves.

c) $\alpha = 0,\ \beta \neq 0$. Analogously the tangent $v-$surfaces along the $u-$curves are of cylinder types. If $\beta_v = 0$ then the tangent $u-$ surfaces along the $v-$ curves are of cone types. In the case when $\beta_v \neq 0$ the tangent $u-$surfaces are determined by the tangents of a curve.

Remark. The relation $\bar{r}_{uv} = \alpha\bar{r}_u$ is a partial differential equation for the unknown coordinate functions $x(u,v)$, $y(u,v)$, $z(u,v)$. If we limite ourselves on the separable product form of these functions we get the class of surfaces $\mathcal{P}_{kv}$ with the coordinate expression:

$$(*) \qquad \bar{r}(u,v) = (f_1(u)g(v),\ f_2(u)g(v),\ g_3(v)),\ \alpha = g_v/g.$$

These surfaces can be created from a plane curve $(f_1(u), f_2(u))$ by homotheties with the coefficient $g(v)$ and with the center in the origin and by translations $g_3(v)\bar{k}$, where $\bar{k}$ is the unit vector of the third axis.
As an example we introduce the sphere S^2 $\bar{r}(u,v) = (r\cos u\cos v, r\sin u\cos v, r\sin v)$

d) $\bar{r}_{uv} = \alpha\bar{r}_u + \beta\bar{r}_v,\ \alpha\cdot\beta \neq 0$.

If $\alpha_u = \beta\alpha$ or $\alpha_u \neq \beta\alpha$ then the tangent $v-$surface along an $u-$curve is a cone surface or a surface of all tangents of a curve respectively.

If $\beta_v = \beta\alpha$ or $\beta_v \neq \beta\alpha$ then the tangent $u-$surface along an $v-$curve is a cone surface or a surface of all tangent of a curve respectively.

If we limite ourselves on the separable product form of coordinate functions of surfaces satisfying the linear partial diferential equation $\bar{r}_{uv} = \alpha\bar{r}_u + \beta\bar{r}_v$ of second order we get the following example

$$\bar{r}(uv) = \left(c_1 exp \left(-\int \frac{\lambda_1 \beta(u)}{\alpha_1(u) - \lambda_1} du \right) exp\lambda_1 \int \alpha_2 \, (v) dv, \right.$$
$$c_2 exp \left(-\int \frac{\lambda_2 \beta(u)}{\alpha_1(u) - \lambda_2} du \right) exp\lambda_2 \int \alpha_2(v) dv,$$
$$\left. c_3 exp \left(-\int \frac{\lambda_3 \beta(u)}{\alpha_1(u) - \lambda_3} du \right) exp\lambda_3 \int \alpha_2 dv \right),$$

where $c_1, c_2, c_3 \lambda_1, \lambda_2, \lambda_3$ are arbitrary constants and $\beta(u), \alpha_1(u), \alpha_2(v)$ are arbitrary functions.

Remark 2. Every line surface $\mathcal{P}$ can be parametrized by the equation

$$\bar{r}\,(u, v) = \bar{a}\,(u) + v\bar{b}\,(u), \tag{6}$$

where $\bar{a}\,(u)$ is the so-called determining curve and $\bar{b}\,(u)$ is the vector of the surface line going cross the point $\bar{a}(u)$. This parametric net we will called natural. Let us recall that this parametrization does not to be suitable from the point of the picture of this surface but can be useful from the applicability point of wiew. It is clear that the tangent $v-$surfaces are identified with $\mathcal{P}$. The tangent $u-$surfaces of the surface (6) are presented by the equation

$$\bar{R}\,(v, t, u_0) = \bar{a}\,(u_0) + v\bar{b}(u_0) + t\left(\bar{a}_u\,(u_0) + v\bar{b}_u\,(u_0)\right).$$

They are developable iff the tangent $v-$surfaces are developable, i.e. iff the surface (6) is developable, i.e. iff $\bar{b}_u(\bar{a}_u \times \bar{b}) = 0$. In this case $\bar{r}_{uv} = \bar{b}_u, \bar{r}_u = \bar{a}_u + v\bar{b}_u, \bar{r}_v = \bar{b}$.

There are two cases:

a) $\bar{b} = c(u)\bar{a}_u$, i.e. the surface (6) is the surface of tangents of the curve $\bar{a}(u)$. Then
$$\bar{b}_u = c_u\bar{a}_u + c\bar{a}_{uu} = \frac{1}{v}\left(\bar{a}_u + v(c_u\bar{a}_u + c\bar{a}_{uu})\right) - \frac{1}{cv}\,c\,\bar{a}_u,$$
i.e. $\alpha = \frac{1}{v}, \beta = -\frac{1}{cv}$. Then $\beta_v = \frac{1}{cv^2} \neq \beta\alpha$, i.e. the tangent $u-$surfaces are also surfaces of tangents of curves.

b) $\bar{b} \neq c\,(u)\,\bar{a}_u$. Then $\bar{b}_u = c_1\bar{a}_u + c_2\bar{b} = \frac{c_1}{1+vc_1}\left(\bar{a}_u + v\,(c_1\bar{a}_u + c_2\bar{b})\right) + c_2 \cdot \frac{1}{1+vc_1}\right)\bar{b}$,

i.j. $\alpha = \frac{c_1}{1+vc_1}, \quad \beta = c_2\frac{1}{1+vc_1}$.

We are looking for a such function $g\,(u)$, the curve $\mathcal{R}\,(u) = \bar{a}\,(u) + g\,(u)\bar{b}(u)$ to be the edge of regression of the surface (6). It satisfies

$$\bar{a}_u + g_u\,\bar{b} + g\,\bar{b}_u = c\bar{b}, \; i.e. \quad \bar{a}_u + g_u\,\bar{b} + g\,(c_1\bar{a}_u + c_2\bar{b}) = c\bar{b}.$$

It is true iff
$$1 + g\,c_1 = 0 \quad , \quad g_u + g\,c_2 = c.$$

If $c_1 = 0$ then the surface (6) is a cylinder. Then $\beta = c_2$. If $c_2 = 0$, i.e. if $\bar{b}_u = \bar{0}$ then the tangent $u-$surfaces are also cylinders. If $c_2(u) \neq 0$ then $\beta_v = c_{2\,v} = 0$ and so the tangent $u-$surfaces are cones.

If $c_1 \neq 0$ and $c = 0$ then the surface (6) is a cone. In this case

$$g = -\frac{1}{c_1}\ ,\ c_2 = \frac{c'_{1u}}{c_1}\ ,\ \beta = \frac{c'_{1u}}{c_1}\,\frac{1}{1+vc_1}\ ,\ \beta_v = -\frac{c'_{1u}}{(1+vc_1)^2}\ ,\ \alpha\beta\frac{c'_{1u}}{(1+vc_1)^2}.$$

We get. If $c'_1 = 0$ then the tangent $u-$surface is a cylinder. If $c'_1 \neq 0$ then the tangent $u-$surface is a surface of tangents of a curve.

If $c_1 \cdot c \neq 0$ then the surface (6) is a surface of tangents of a curve. Then $\alpha\beta \neq \beta_v$ and so the tangent $u-$surfaces are also surfaces of tangents of curves.

II. If $\bar{r}_{uv}, \bar{r}_u, \bar{r}_v$ are linearly independent then both the tangent $v-$ and $u-$ surfaces of the surface $\mathcal{P}$ are undevelopable. If the surface $\mathcal{P}$ is a line surface then it has this property if and only if is undevelopable.

Remark 3. (About graphs of the functions $z = f(x,y)$ of two variable.) Let

$$\bar{r}\,(u,v) = (u, v, f(u,v))$$

be the natural parametrization of the surface given by a function $z = f(x,y)$. In this case $\bar{r}_{uv} = (0,0,f_{uv}), \bar{r}_u = (1,0,f_u), \bar{r}_v = (0,1,f_v)$. Therefore both the tangent $u-$ and $v-$ surfaces along those parametric curves are developable iff $f_{uv} = 0$. In this case $\alpha = 0 = \beta$ and both the tangent $u-$ and $v-$surfaces are cylinders. If $f_{uv} \neq 0$ then the tangent surfaces are undevelopable.

Parametrization of the surfaces given by an equation $F(x,y,z) = 0$

Let a surface $\mathcal{P}$ is given by an equation $F(x,y,z) = 0$ and let $\bar{r}\,(u,v) = (x(u,v),\ y\,(u,v),\ z\,(u,v))$ be its parametric representation, i.e. let the equation

$$F\ (x\,(u,v),\ y\,(u,v), z\,(u,v)) = 0 \tag{7}$$

is the identity for $(u,v) \in \Omega$. Derivating the identity (7) with respect to u and v we get

$$F_x\, x_u + F_y y_u + F_z z_u = 0\ ,\ F_x x_v + F_y y_v + F_z z_v = 0. \tag{8}$$

Then the derivative of the first part of (8) with respect to v gives

$$F_{xx} x_u x_v + F_{xy}(x_u y_v + x_v y_u) + F_{xz}(x_u z_v + x_v z_u) + F_{yy} y_u y_v + F_{yz}(y_u z_v + z_u y_v) +$$
$$+ f_{zz} z_u z_v + F_x x_{uv} + F_y y_{uv} + F_z z_{uv} = 0. \tag{9}$$

The equation (8) and (9) can by shortly rewrite in the forms

$$dF\,(\bar{r}_u) = 0\,, dF\,(\bar{r}_v) = 0\,, d^2F\,(\bar{r}_u, \bar{r}_v) + dF\,(\bar{r}_{uv}) = 0.$$

Proposition 2. *Let $\bar{r}(u,v)$ be a parametric representation of the surface $\mathcal{P}$ given by the equation $F(x,y,z)=0$. Then the tangent $v-$ and $u-$ surfaces are developable iff $d^2F(\bar{r}_u,\bar{r}_v)=0$, i.e. iff the tangent vector $\bar{r}_u,\bar{r}_v$ at any point of $\mathcal{P}$ vanish the differential d^2F of second order.*

Proof. If $\bar{r}(u,v)$ is a parametric representation of $\mathcal{P}$ then the equation (8) and (9) are satisfied. Let $d^2F(\bar{r}_u,\bar{r}_v)=0$. Then the relation (9) gives $dF(\bar{r}_{uv})=0$, i.e. $\bar{r}_{uv}$ is a tangent vector of the surface $\mathcal{P}$, i.e. $\bar{r}_{uv}=\alpha\bar{r}_u+\beta\bar{r}_v$. Conversaly if $\bar{r}_{uv}=\alpha\bar{r}_u+\beta\bar{r}_v$ then $dF(\bar{r}_{uv})=0$ and then the relation (9) completes our proof.

Remark 4. It is easy to see that a parametric representation of the sphere $x^2+y^2+z^2=r^2$ has the tangent $v-$ and $u-$ surfaces which are developable if and only if it is orthogonal.

Proposition 3. *Let a surface $\mathcal{P}$ is given by the equation $F(x,y,z)=0$. Then there exists a such parametric representation of the surface $\mathcal{P}$ that its tangent $v-$ and $u-$ surfaces are developable if and only if there exist such two vector fields X_1,X_2 which at points of $\mathcal{P}$ satisfy the equations $dF(X_1)=0,\ dF(X_2)=0,\ d^2F(X_1,X_2)=0$ and $[X_1,X_2]$=0, where $[X_1,X_2]$ is the Lie bracket of the vector fields X_1,X_2.*

Proof. The necessary condition is clear. It is well known that if $[X_1,X_2]=0$ then there exists a such parametric representation $\bar{r}=\bar{r}(u,v)$ that $\bar{r}_u=X_1$, $\bar{r}_v=X_2$, see [3]. If $dF(X_1)=0$, $dF(X_2)=0$ then $\bar{r}=\bar{r}(u,v)$ is a parametric representation of $\mathcal{P}$. If $d^2F(X_1,X_2)=0$ then by Proposition 2 the tangent of $v-$surfaces are developable.

Remark 5. From the graphic point of view the most suitable are the parametric representation the tangent $v-$ and $u-$ surfaces of which are cylinders or cones. So we prefer the parametric representations of the classes Ia, Ib, Ic.

Example. Consider the sphere $x^2+y^2+z^2=r^2$. Look for a parametric representation of the type $(*)$ from the class Ib, $\bar{r}(u,v)=(f_1(u)g(v), f_2(u)g(v), g_3(v))$. Using the relations (8) and (9) it is easy to infer the following conditions for the functions $f_1(u), f_2(u)$, $g(v)$, $g_3(v)$: $f_1^2+f_2^2=k^2$, k is constant

$$k^2g^2+g_3^2=r^2.$$

For $d=r>0, f_1=r\cos u$, $f_2=r\sin u$, $g=\cos v$, $g_3=r\sin v$ we get the spherical representation of the consider sphere which is a global parametric representation with two singular isolated points.

References

[1] Budinský, B., Kepr, B., *Základy diferenciální geometrie s technickými aplikáciami*, SNTL Praha, 1970.

[2] Guggenheimer, H. W., *Differential geometry*, Dover publications, INC, New York, 1977.

[3] Kolář, I., Michor, P.W., Slovák, J., *Natural operations in differential geometry*, Springer-Verlag, 1993.

(Received July 10, 2000)

Dept. of Computer Science
Matej Bel University
Tajovského 40
974 01 Banská Bystrica
SLOVAKIA

Dept. of Mathematics
TU Zvolen
Masarykova 24
960 53 Zvolen
SLOVAKIA

E-mail address: dekret@fpv.umb.sk
baksa@vsld.tuzvo.sk

Acta Univ. M. Belii
Math. no. 8(2000), pp. 11-19

THE LATTICE OF VARIETIES OF GRAPHS

Alfonz Haviar

Abstract. In the paper we investigate classes of graphs closed under isomorphic images, subgraph identifications and contractions and we study the lattice of these classes.

0. Introduction

By a graph $\mathcal{G} = (V, E)$ we mean an undirected connected finite graph without loops and multiple edges. We denote the set of all vertices of a graph $\mathcal{G}$ by $V(\mathcal{G})$ and the set of all edges by $E(\mathcal{G})$. An edge $\{u, v\}$ is briefly denoted by uv. We denote the complete n-vertices graph by $\mathcal{K}_n$ and the n-vertices circle (in which every vertex is of degree two) by $\mathcal{C}_n$.

A class of all graphs closed under isomorphic images is called a property of graphs (for example in [1]) or a variety of graphs (in [5]). To put considerations in the right context within set theory, we will assume that the vertex sets of all considered graphs are subsets of a fixed countable infinite set W, and we talk about graphs *over* W.

The set of all varieties of graphs for which vertex sets are subsets of W with set inclusion as the partial ordering is a complete lattice isomorphic to the Boolean lattice $P(W)$ of all subsets of the set W. The atoms of this lattice are the varieties which are generated generated by only one graph. In theory of graphs we are interested in varieties of graphs closed under more closed operators, for example varieties closed under induced subgraphs [11], varieties closed under induced subgraphs and identifications [5], varieties closed under generalized hereditary operators [1], [2], [9], etc.

One of the most important operators in theory of graphs is the operator of contraction (of edges). It produces "smaller" graphs. A natural operator producing "bigger" graphs is the operator of identification in (connected) induced subgraphs. In this paper we pay attention to varieties of graphs closed under identifications and contractions.

A set of all varieties of graphs closed under given closure operators with set inclusion as the partial ordering is a complete lattice ([3], Theorem 5.2, p. 18]). The smallest variety containing a set $\mathbb{K}$ of graphs is denoted by $V(\mathbb{K})$ and we call it the variety generated by $\mathbb{K}$. If $\mathbb{K} = \{\mathcal{G}_1, \ldots, \mathcal{G}_n\}$ we simply denote it by $V(\mathcal{G}_1, \ldots, \mathcal{G}_n)$.

2000 *Mathematics Subject Classification.* 05C99.
Key words and phrases. Graph, subgraph identification, contraction, variety.
This research is supported by the Ministry of Education of Slovak republic, grant No. 1/6132/99.

1. Preliminary results.

Our aim in this paper is to investigate varieties of graphs closed under suborgraph identifications and contractions. The following operation of *a subgraph identification* of graphs in a connected induced subgraph generalizes the operation of the union of graphs and was introduced in [5].

Definition 1.1. Let $\mathcal{G}_1 = (V_1, E_1)$ and $\mathcal{G}_2 = (V_2, E_2)$ be disjoint graphs. Let $\mathcal{G}_1' = (V_1', E_1')$ and $\mathcal{G}_2' = (V_2', E_2')$ be connected induced subgraphs of $\mathcal{G}_1$, $\mathcal{G}_2$, respectively and let $f : G_1' \to G_2'$ be an isomorphism. The subgraph identification of $\mathcal{G}_1$ with $\mathcal{G}_2$ under f is the graph $\mathcal{G} = \mathcal{G}_1 \cup^f \mathcal{G}_2 = (V, E)$, where

$$V = V_1 \cup (V_2 - V_2'),$$
$$E = \{uv \mid u, v \in V \quad \text{and} \quad uv \in E_1 \cup E_2 \text{ or } f(u)v \in E_2\}.$$

If graphs $\mathcal{G}_1$ and $\mathcal{G}_2$ are not disjoint we may take instead of the graph $\mathcal{G}_2$ a graph $\mathcal{G}_3$ isomorphic with $\mathcal{G}_2$ and disjoint with $\mathcal{G}_1$ (for details see [5]). When no confusion can arise we will simply talk about the subgraph identification under the induced subgraph $\mathcal{G}_1'$ or about gluing in the induced subgraph $\mathcal{G}_1'$.

The fact that $f : G_1' \to G_2'$ is an isomorphism of the connected induced subgraph $\mathcal{G}_1' \subseteq \mathcal{G}_1$ onto the connected induced subgraph $\mathcal{G}_2' \subseteq \mathcal{G}_2$ will be denoted by $f : G_1 \rightarrowtail G_2$.

It is easy to see that $\mathcal{G}_1 \cup^f \mathcal{G}_2 \cong \mathcal{G}_2 \cup^{f^{-1}} \mathcal{G}_1$ and if f is an authomorphism of a graph $\mathcal{G}$ then $\mathcal{G} \cup^f \mathcal{G} = \mathcal{G}$. Clearly, a subgraph identification of connected graphs is again a connected graph.

Lemma 1.1 (see [5]). *Let $\mathcal{G} = (V, E)$ be a connected graph, which is neither a complete graph nor a circle. Then there are two nonadjacent vertices $u, v \in V(G)$ such that $\mathcal{G} - \{u, v\}$ is a connected graph.*

Corollary 1.2. *If $\mathcal{G}$ is a graph which is neither a circle nor a complete graph, then $\mathcal{G}$ contains proper connected induced subgraphs $\mathcal{G}_1$, $\mathcal{G}_2$ such that $\mathcal{G} \cong \mathcal{G}_1 \cup^f \mathcal{G}_2$, where $f : G_1 \rightarrowtail G_2$.*

Definition 1.2 ([4]). We say that a graph $\mathcal{G}_2$ is a contraction of a graph $\mathcal{G}_1$ if there exists a one-to-one correspondence between $V(G_2)$ and the elements of a partition of $V(G_1)$ such that each element of the partition induces a connected subgraph of $\mathcal{G}_1$, and two vertices of $\mathcal{G}_2$ are adjacent if and only if the subgraph induced by the union of the corresponding subsets is connected.

If adjacent vertices $u, v \in V(G_1)$ belong to the same block of the partition of the set $V(G_1)$ we will say that the vertices u, v have been identifying by the contraction.

If a graph $\mathcal{G}_2$ is a contraction of a graph $\mathcal{G}_1$, we write $\mathcal{G}_2 \lhd \mathcal{G}_1$.

Let $\mathbb{K}$ be a family of graphs. Denote

$$\gamma(\mathbb{K}) = \{\mathcal{G}_1 \cup^f \mathcal{G}_2; \mathcal{G}_1, \mathcal{G}_2 \in \mathbb{K}, \quad f : G_1 \rightarrowtail G_2\},$$

$$C(\mathbb{K}) = \{\mathcal{G} : \mathcal{G} \lhd \mathcal{G}' \text{ for some graph } \mathcal{G}' \in \mathbb{K}\},$$

$I(\mathbb{K})$ – the set of all isomorphic images of graphs in $\mathbb{K}$.

Since $G \cup^{id} G = G$ and $\mathcal{G} \lhd \mathcal{G}$ we have
$\mathbb{K} \subseteq \gamma(\mathbb{K}) \subseteq \gamma^2(\mathbb{K}) \subseteq \cdots \subseteq \gamma^n(\mathbb{K}) \subseteq \ldots,$
$\mathbb{K} \subseteq C(\mathbb{K}) \subseteq C^2(\mathbb{K}) \subseteq \cdots \subseteq C^n(\mathbb{K}) \subseteq \ldots,$
for any set $\mathbb{K}$ of graphs. Note that $O^n(\mathbb{K}) = O^{n-1}(O(\mathbb{K}))$, for each $n > 1$.

Definition 1.3. A set $\mathbb{K}$ of graphs over W is said to be a variety of graphs closed under subgraph identifications and contractions if

$$I(\mathbb{K}) \subseteq \mathbb{K} \quad \& \quad \gamma(\mathbb{K}) \subseteq \mathbb{K} \quad \& \quad C(\mathbb{K}) \subseteq \mathbb{K}.$$

It is obvious that the operators C and γ are closure operators on the system of all sets of graphs over W. Thus, the next statement holds.

Proposition 1.3. *The set of all varieties of graphs over W closed under subgraph identifications and contractions with the set inclusion as the partial ordering is a complete lattice.*

Let $\mathbb{K}$ be a set of graphs. Define the operator σ by
$\sigma(\mathbb{K}) = (C\gamma)(\mathbb{K}) \cup (C\gamma)^2(\mathbb{K}) \cup \cdots = \cup_{n=1}^{\infty}(C\gamma)^n(\mathbb{K}),$
where $(C\gamma)(\mathbb{K}) = C(\gamma(\mathbb{K}))$ and $(C\gamma)^n(\mathbb{K}) = C(\gamma((C\gamma)^{n-1}(\mathbb{K}))$ if $n > 1$.

Theorem 1.4. *For every set $\mathbb{K}$ of graphs*

$$V(\mathbb{K}) = \sigma(\mathbb{K}).$$

Proof. Let $\mathcal{G}_1, \mathcal{G}_2 \in \sigma(\mathbb{K})$ and let $f : G_1 \rightarrowtail G_2$. Then there exist m, n such that $\mathcal{G}_1 \in (C\gamma)^n(\mathbb{K})$ and $\mathcal{G}_2 \in (C\gamma)^m(\mathbb{K})$. We see at once that $n < m$ implies $\mathcal{G}_1, \mathcal{G}_2 \in (C\gamma)^m(\mathbb{K})$ and so $\mathcal{G}_1 \cup^f \mathcal{G}_2 \in (C\gamma)^{m+1}(\mathbb{K})$. Similarly, $\mathcal{G} \in (C\gamma)^n(\mathbb{K})$ and $\mathcal{G}_1 \lhd \mathcal{G}$ yields $\mathcal{G}_1 \in (C\gamma)^{n+1}(\mathbb{K})$. Thus, we have shown that $\sigma(\mathbb{K})$ is a variety of graphs closed under subgraph identification and contraction and it contains the set $\mathbb{K}$. Consequently $V(\mathbb{K}) \subseteq \sigma(\mathbb{K})$. The opposite inclusion is obvious. □

2. The lattice of varieties of graphs

In this section we investigate the lattice of all varieties of graphs closed under identifications and contractions. This lattice is denoted by $\mathcal{L}$.

Clearly, the least element of the lattice $\mathcal{L}$ is the vatiety $V(\mathcal{K}_1)$, where $\mathcal{K}_1$ is a one-vertex graph. We will denote it by $\mathbf{0}$.

Proposition 2.1. *The variety $V(\mathcal{K}_2)$ generated by the two-vertex graph is the variety of all trees. Moreover, it is the only atom of the lattice $\mathcal{L}$.*

Proof. Let $\mathbb{V} \neq \mathbf{0}$ be an element of the lattice $\mathcal{L}$ and let $\mathcal{G} \in \mathbb{V}$, $\mathcal{G} \neq \mathcal{K}_1$. It is easy to see that the graph $\mathcal{K}_2$ is a contraction of $\mathcal{G}$ and so $\mathcal{K}_2 \in \mathbb{V}$, which implies $V(K_2) \subseteq \mathbb{V}$. Using an induction on a number of vertices we see that every tree belongs to $V(K_2)$. On the other hand, no graph $\mathcal{G}$ in $V(K_2)$ contains a circle (a contraction of a tree is again a tree and a subgraph identification of trees is a tree, too).

Lemma 2.2. *The only variety covering the variety of all trees in $\mathcal{L}$ is the variety $V(\mathcal{C}_3)$.*

Proof. Let $\mathbb{V}$ be an element of $\mathcal{L}$ for which $V(\mathcal{K}_2) < \mathbb{V}$. The variety $\mathbb{V}$ contains a graph $\mathcal{G}$ containing a circle $\mathcal{C}_n$. This clearly forces $\mathcal{C}_3 \in \mathbb{V}$. (It is obvious that $\mathcal{C}_3$ is a contraction of the circle $\mathcal{C}_n$ and so $\mathcal{C}_3 \lhd \mathcal{G}$.) Therefore $V(\mathcal{C}_3) \subseteq \mathbb{V}$. □

Lemma 2.3. *Let $\mathcal{G}$ be a graph belonging to a variety $\mathbb{V} \geq V(\mathcal{C}_3)$. If the set $E(\mathcal{G})$ contains edges uv and uw but it does not contain the edge vw then the variety $\mathbb{V}$ also contains the graph $\tilde{\mathcal{G}}$ given by*

$V(\tilde{\mathcal{G}}) = V(\mathcal{G})$ and $E(\tilde{\mathcal{G}}) = E(\mathcal{G}) \cup \{vw\}$.

Proof. By Lemmas 2.1 and 2.2 the variety $\mathbb{V}$ contains the graph $\mathcal{C}_3$, hence it contains the identification $\mathcal{H}$ of two copies of $\mathcal{C}_3$, where $V(\mathcal{H}) = \{u'.v', w', x\}$ and $E(\mathcal{H}) = \{u'v', u'w', u'x, v'w', w'x\}$ (see Fig. 1). Let $f : G \rightarrowtail H$ be given by

$f(u) = u'$, $f(v) = v'$ and $f(w) = x$.

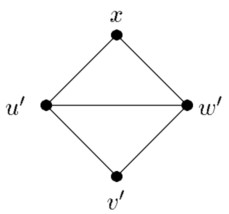

Fig. 1

Now, the graph $\tilde{\mathcal{G}}$ is obtained by contraction of the edge ww' of the graph $\mathcal{G} \cup^f \mathcal{H}$. This yields $\tilde{\mathcal{G}} \in \mathbb{V}$.

Corollary 2.4. *The variety $V(\mathcal{C}_3)$ contains all complete graphs.*

Definition 2.1. We will say that a graph $\tilde{\mathcal{G}}$ is a triangular cover of a graph $\mathcal{G}$ if $\tilde{\mathcal{G}}$ can be obtained from $\mathcal{G}$ by adding edges as in Lemma 2.3.

Theorem 2.5. *If a circle $\mathcal{C}_m$ is a contraction of an identification $\mathcal{G}_1 \cup^f \mathcal{G}_2$ then $\mathcal{C}_m$ is a contraction of a graph $\tilde{\mathcal{G}}_1$ or $\tilde{\mathcal{G}}_2$, where $\tilde{\mathcal{G}}_i$ is a suitable triangular cover of the graph $\mathcal{G}_i$, $i \in \{1, 2\}$.*

Proof. Let $\mathcal{C}_m \lhd \mathcal{G}_1 \cup^f \mathcal{G}_2$, and let $\{A_1, A_2, \ldots, A_m\}$ be a partition of the set $V(\mathcal{G}_1 \cup^f \mathcal{G}_2)$ corresponding to the above contraction of $\mathcal{G}_1 \cup^f \mathcal{G}_2$ to $\mathcal{C}_m$.

a) Let there exist a block A_i disjoint with the set $V(\mathcal{G}_1')$ and let $A_i \subseteq V(\mathcal{G}_1)$. Let $A_j \cap V(\mathcal{G}_1') \neq \emptyset$, $A_k \cap V(\mathcal{G}_1') \neq \emptyset$, $1 \leq j \leq i \leq k \leq m$ and $A_q \cap V(\mathcal{G}_1') = \emptyset$ for each $q \in \{j+1, \ldots, k-1\}$ (recall that the graphs are glued in the subgraph $\mathcal{G}_1'$). For vertices $w_1 \in A_j \cap V(\mathcal{G}_1')$, $w_2 \in A_k \cap V(\mathcal{G}_1')$, there is a path

(p_1) $w_1 = v_0, v_1, \ldots, v_r = w_2$

in the subgraph $\mathcal{G}_1'$, therefore $A_p \cap V(\mathcal{G}_1') \neq \emptyset$ for each $p \in \{1, \ldots, j, k, \ldots, m\}$. There exist also a path

(p_2) $w_2 = u_0, u_1, \ldots, u_s = w_1$

in the graph $\mathcal{G}_1$ disjoint with (p_1), i.e. there exists a circle $\mathcal{C}$ in the graph $\mathcal{G}_1$ for which $V(\mathcal{C}) \cap A_q \neq \emptyset$ for each $q \in \{1, \ldots, m\}$. Denote

$B_1 = A_1 \cap V(\mathcal{G}_1), \quad \ldots, \quad B_m = A_m \cap V(\mathcal{G}_1)$.
If the subgraphs of the graph $\mathcal{G}_1$ induced by the sets $B_1, \ldots, B_m$ are connected then $\mathcal{C}_m \lhd \mathcal{G}_1$.

Let there exist (for example) vertices $x \in B_1$, $y \in B_1 \cap V(\mathcal{C})$ for which there is no path from x to y in the subgraph of $\mathcal{G}_1$ induced by B_1. Denote by
$x = z_0, z_1, \ldots, z_t = y$
a path from x to y in the graph $\mathcal{G}_1$. If the distance of vertices z_p, z_q or z_p, v, $p, q \in \{0, 1, \ldots, t\}$, $v \in V(\mathcal{C})$ is two and these vertices belong to the same block B_l, $l \in \{1, \ldots, m\}$, or belong to adjacent blocks B_l, B_{l+1}, $l \in \{1, \ldots m-1\}$, then we can add the edge $z_p z_q$ and $z_p v$ to a vertex set obtained from $E(\mathcal{G}_1)$ (by Lemma 2.3). After finitely many steps we obtain a graph $\mathcal{G}_1^\star$ such that in the subgraph of $\mathcal{G}_1^\star$ induced by the set B_1 there is a path from x to y. Repeating this proces we can obtain a triangular cover $\tilde{\mathcal{G}}_1$ of the graph $\mathcal{G}_1$ such that the subgraphs of $\tilde{\mathcal{G}}_1$ induced by the sets $B_1, \ldots, B_m$ are connected and so $\mathcal{C}_m \lhd \tilde{\mathcal{G}}_1$.

b) Let $A_q \cap V(\mathcal{G}_1') \neq \emptyset$ for each $q \in \{1, \ldots, m\}$ and let
$v_1 \in A_1 \cap V(\mathcal{G}_1'), \ldots, v_m \in A_m \cap V(\mathcal{G}_1')$.
There exists a path from v_l to v_{l+1} in the graph $\mathcal{G}_1'$ for each $l \in \{1, \ldots, m\}$ (we compute modulo m). Hence there exists a circle $\mathcal{C}$ of the graph $\mathcal{G}_1$ or of the graph induced by the set $V(\mathcal{G}_1') \cup (V(\mathcal{G}_2) - V(\mathcal{G}_2'))$ (the natural copy of the graph $\mathcal{G}_2$) in the graph $\mathcal{G}_1 \cup^f \mathcal{G}_2$ for which $V(\mathcal{C}) \cap A_q \neq \emptyset$ for each $q \in \{1, \ldots, m\}$. Thus, the next part of the proof runs in the same way as the above corresponding part of the proof.

Theorem 2.6. *If a graph $\mathcal{G}$ belongs to $V(\mathcal{C}_m)$ then with each circle $\mathcal{C}$ of $\mathcal{G}$ contains a plane subgraph with the exterior face $\mathcal{C}$ and regions $\mathcal{C}_n$, $3 \leq n \leq m$.*

Proof. Let $\mathcal{G} \in V(\mathcal{C}_m)$.

a) The statement holds if $\mathcal{G} = \mathcal{C}_m$.

b) Let graphs $\mathcal{G}_1, \mathcal{G}_2 \in V(\mathcal{C}_m)$ contain with each circle $\mathcal{C}$ also a plane subgraph with the exterior $\mathcal{C}$ and regions $\mathcal{C}_n$, $3 \leq n \leq m$ and let $\mathcal{C} = v_1 v_2 \ldots v_n v_1$ be a circle of $\mathcal{G}_1 \cup^f \mathcal{G}_2$. If $\mathcal{C}$ is a subgraph of $\mathcal{G}_1$ or $\mathcal{G}_2$ then $\mathcal{G}_1$ or $\mathcal{G}_2$ and so also $\mathcal{G}$ contains a plane subgraph with the exterior $\mathcal{C}$ and regions $\mathcal{C}_n$, $3 \leq n \leq m$. Let $\mathcal{C} = v_1 v_2 \ldots v_n v_1$ be a subgraph neither $\mathcal{G}_1$ nor $\mathcal{G}_2$. Let $v_i, v_{i+1}, \ldots, v_{i+j} \in V(\mathcal{C})$; we will say that $v_i \curvearrowright v_{i+j}$ is a *jump* in $\mathcal{G}$ if $v_i \in V(\mathcal{G}_1) - V(\mathcal{G}_1')$, $v_{i+1}, \ldots, v_{i+j-1} \in V(\mathcal{G}_1')$, $v_{i+j} \in V(\mathcal{G}_2) - V(\mathcal{G}_2')$ or $v_i \in V(\mathcal{G}_2) - V(\mathcal{G}_2')$, $v_{i+1}, \ldots, v_{i+j-1} \in V(\mathcal{G}_1')$, $v_{i+j} \in V(\mathcal{G}_1) - V(\mathcal{G}_1')$ (see Fig. 2).

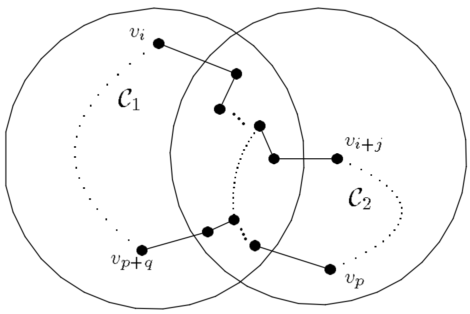

Fig. 2

We proceed by induction on the number of jumps of the circle $\mathcal{C}$. Firstly, we suppose that there are only two jumps in $\mathcal{G}_1 \cup^f \mathcal{G}_2$, $v_i \curvearrowright v_{i+j}$ and $v_p \curvearrowright v_{p+q}$, $i < p$. Since $\mathcal{G}_1 \cup^f \mathcal{G}_2$ is the subgraph identification under a connected subgraph $\mathcal{G}_1'$, there exists a path $v_{i+1}, w_1, w_2, \ldots, w_k, v_{p+q-1}$ in $\mathcal{G}_1'$. If this path is disjoint with the circle $\mathcal{C}$, we get a circle $\mathcal{C}^{(1)}$ of the graph $\mathcal{G}_1$ and a circle $\mathcal{C}^{(2)}$ of the graph $\mathcal{G}_2$ which both contain the path $(v_{i+1}, w_1, w_2, \ldots, w_k, v_{p+q-1})$ or its part (see Fig. 2). By assumptions there exist plane subgraphs with exteriors faces $\mathcal{C}^{(1)}$ and $\mathcal{C}^{(2)}$ and regions $\mathcal{C}_n$, $3 \le n \le m$. If the mentioned path is not disjoint with the circle $\mathcal{C}$ we get circles $\mathcal{C}^{(1)}, \ldots, \mathcal{C}^{(k)}$ such that there exist plane subgraphs with exteriors faces $\mathcal{C}^{(1)}, \ldots, \mathcal{C}^{(k)}$ and regions $\mathcal{C}_n$, $3 \le n \le m$, each of them belongs to either $\mathcal{G}_1$ or $\mathcal{G}_2$ and each of them contains a part of the path $(v_{i+1}, w_1, w_2, \ldots, w_k, v_{p+q-1})$ (see Fig. 3).

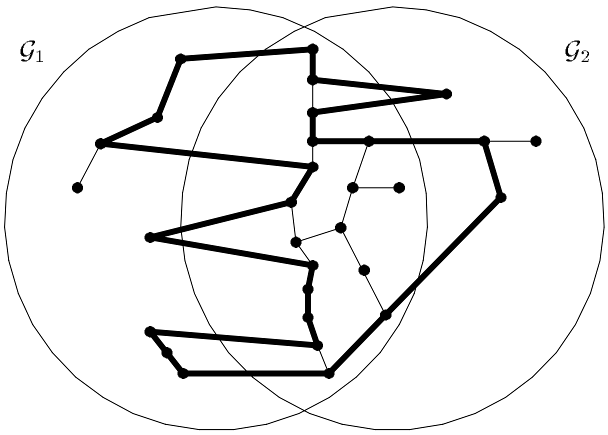

Fig. 3

Since one can get the circle $\mathcal{C}$ by successive gluing the circles $\mathcal{C}^{(1)}, \ldots, \mathcal{C}^{(k)}$, there exists plane subgraph with exterior face $\mathcal{C}$ and regions $\mathcal{C}_n$, $3 \le n \le m$, too. Assuming the statement for circles with less than $2r$ jumps, we will prove it for $2r$ jumps. Without loss of generality we can assume that $v_i \curvearrowright v_{i+j}$, $v_p \curvearrowright v_{p+q}$ are jumps and that for each jump $v_k \curvearrowright v_{k+s}$ of the circle $\mathcal{C}$, $i \le k \le p$ holds. Analogously as in the case of two jumps we can get circles $\mathcal{C}^{(1)}, \ldots, \mathcal{C}^{(k)}$ having less than $2r$ jumps. By assumption there exist plane subgraphs with exteriors faces $\mathcal{C}^{(1)}, \ldots, \mathcal{C}^{(k)}$ and regions $\mathcal{C}_n$, $3 \le n \le m$, therefore there exists plane subgraph with exterior face $\mathcal{C}$ and regions $\mathcal{C}_n$, $3 \le n \le m$, too.

c) Let $\mathcal{G} \triangleleft \mathcal{G}'$, $\mathcal{G}' \in V(\mathcal{C}_m)$ and let $\mathcal{C} = v_1 v_2 \ldots v_n v_1$ be a circle of $\mathcal{G}$. Let $A_1, A_2, \ldots, A_s$ be a partition of $V(G')$ corresponding to the contraction of $\mathcal{G}'$ to $\mathcal{G}$. Without loss of generality we can assume that

$v_1 \in A_1$, …, $v_n \in A_n$.

For any blocks A_i, A_{i+1}, $1 \le i \le n$ (we compute modulo n) there are vertices $w_i \in A_i$ and $w_{i+1} \in A_{i+1}$ for which $w_i w_{i+1} \in E(G')$. The subgraphs induced by sets A_i and A_{i+1} are connected, hence there exists a path from v_i to v_{i+1} in $\mathcal{G}'$. It implies that there exists (in $\mathcal{G}'$) a circle $\mathcal{C}'$ with vertices from A_1, … , A_n having a contraction the circle $\mathcal{C}$. By assumption there is a plane subgraph of $\mathcal{G}'$ with exterior face $\mathcal{C}'$ and regions C_n, $3 \leq n \leq m$. and by contracting edges we can get from it the plane subgraph with exterior face $\mathcal{C}$ and regions C_n, $3 \leq n \leq m$.

Corollary 2.7. *The lattice $\mathcal{L}$ contains the infinite chain*
$\mathbf{0} < V(\mathcal{K}_2) < V(\mathcal{C}_3) < V(\mathcal{C}_4) < \cdots < V(\mathcal{C}_n) < \cdots < \mathbf{1}$
where the variety $\mathbf{1}$ *is generated by the set of all circles.*

Proof. By Theorem 2.6 we have $\mathcal{C}_{n+1} \notin V(\mathcal{C}_n)$ for each $n \geq 3$. It follows from Corollary 1.2 that the variety $\mathbf{1}$ is the greatest element of $\mathcal{L}$.

Theorem 2.8. *The variety $V(\mathcal{C}_4)$ does not cover the variety $V(\mathcal{C}_3)$ and the variety $V(\mathcal{C}_5)$ does not cover the variety $V(\mathcal{C}_4)$*

Proof. Let us denote by $\mathcal{G}_{3-4}$ the graph in Fig. 4

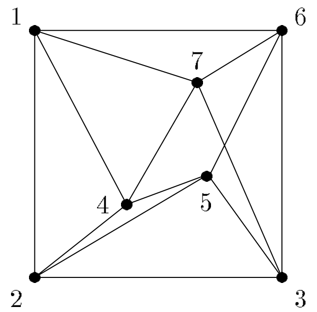

Fig. 4

It is obvious that $V(\mathcal{C}_3) \leq V(\mathcal{G}_{3-4}) \leq V(\mathcal{C}_4)$. We can check that a plane subgraph of $\mathcal{G}_{3-4}$ with the exterior face $C = (1, 2, 3, 6)$ and regions $\mathcal{C}_3$ does not exist, therefore $\mathcal{G}_{3-4} \notin V(\mathcal{C}_3)$. On the other hand it can be checked that
a) if we add any edge to $E(G_{3-4})$ or
b) make any contraction of the graph $\mathcal{G}_{3-4}$,
we obtain a graph belonging to the variety $V(\mathcal{C}_3)$. Hence $\mathcal{C}_4 \notin V(G_{3-4})$. It implies $V(\mathcal{C}_3) < V(\mathcal{G}_{3-4}) < V(\mathcal{C}_4)$.

We can analogously prove that $V(\mathcal{C}_4) < V(\mathcal{G}_{4-5}) < V(\mathcal{C}_5)$, where $\mathcal{G}_{4-5}$ is the graph in Fig. 5a.

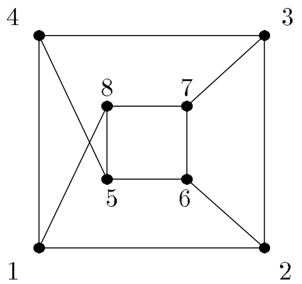

Fig. 5a

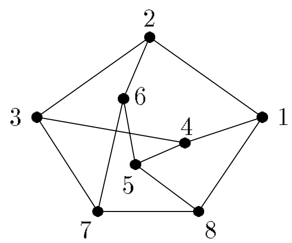

Fig. 5b

Note that the graph depicted in Fig. 5a is depicted in Fig. 5b, too.

The graphs in Fig. 4 and 5 indicate that the structure of the lattice of varieties is not trivial. We will give some problems referring to the lattice $\mathcal{L}$ of varieties.

1. What is the width of the lattice $\mathcal{L}$? (By results of Robertson and Seymour [10], $\mathcal{L}$ does not contain an infinite antichain.)
2. How many varieties cover the variety $V(\mathcal{C}_3)$?
3. What is the length of the interval $[V(\mathcal{C}_3), V(\mathcal{C}_4)]$?
4. Assume $\mathcal{H}$ is the graph in Fig. 6. Is the variety $V(\mathcal{H})$ noncomparable with the variety $V(\mathcal{C}_5)$?

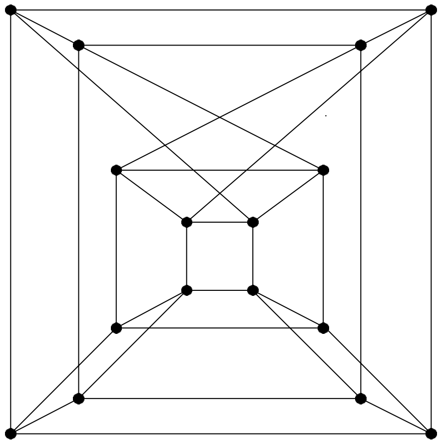

Fig. 6

References

[1] M. Borowiecki, M. Frick, P. Mihók and G. Semanišin, *A survey of hereditary properties of graphs*, Discusiones Math. Graph Theory **17** (1997), 5 - 50.

[2] M. Borowiecki and P. Mihók, *Hereditary properties of graphs*, in: V. R. Kulli,ed., Advances in Graph Theory (Vishwa International Publication, Gulbarga) (1991), 41 - 68.

[3] S. Burris and H. P. Sankappanavar, *A Course in Universal Algebra*, Springer-Verlag, New York Heidelberg Berlin, 1981.

[4] G. Chartrand and L. Lesniak, *Graphs & Digraphs (third ed.)*, Chapman & Hall, London, 1996.

[5] A. Haviar and R. Nedela, *On varieties of graphs*, Discussiones Math., Graph Theory. **18** (1998), 209-223.

[6] R. P. Jones, *Hereditary properties and P-chromatic numbers* (1974), Cambridge Univ. Press,, Cambridge, 83-88.

[7] S. Klaužar and M. Petkovšek, *Notes on hereditary classes of graphs* **25** (1987), Preprint Ser. Dept. Math. University E.K., Ljubljana, no.206.

[8] P. Mihók, *On graphs critical with respect to generalized independence numbers*, Colloquia Mathematica Societatis János Bolyai 52, Combinatorics **2** (1987), 417 - 421.

[9] P. Mihók and R. Vasky, *Hierarchical Decompositions of Diagrams in Information System Analysis and Lattices of Hereditary Properties of Graphs* (2000), Proceedings of ISCM Herlany 1999, ed. A. Haščák, V. Pirč, V Šoltés, University of Technology, Košice, 126-129.

[10] N. Robertson, P. Seymour, *Graph Structure Theory* (1993), Contemporary Mathematics 147, AMS, Providence, Rhode Island, 688 pp.

[11] E. R. Scheinerman, *On the structure of hereditary classes of graphs*, Jour. Graph Theory **10** (1986), 545 - 551.

(Received October 30, 2000)

Dept. of Mathematics
Matej Bel University
Tajovského 40
974 01 Banská Bystrica
SLOVAKIA

E-mail address: haviar@fpv.umb.sk

Acta Univ. M. Belii
Math. no. 8(2000), pp. 21-26

FUZZY DIVERGENCE MEASURES

Inés Couso, Vladimír Janiš, Susana Montes

Abstract. A divergence measure is a tool than can be used to measure how two fuzzy sets differ from each other. Particularly it can be used to estimate the fuzziness measure of a fuzzy set. The existing concepts of a divergence measure do not reflect on what membership values the given fuzzy sets are different. We use a divergence measure whose output is not a real number but a fuzzy quantity and show that this quantity is able to distinguish those pairs of fuzzy sets, for which the classical divergence measure gives identical results.

In recent years several attempts to compare pairs of fuzzy sets have been done either measuring their similarity ([2], [9], [11]) or difference between them ([1]). In relation with the latter paper, Montes et. al ([8]) introduced the definition of a divergence measure. This concept generalizes, except for the symmetry property (that could be excluded from the set of axiom in some particular cases) the concept of dissimilarity measures proposed by Bouchon-Meunier et al. in [1].

Assigning a real number as the value of the difference between two fuzzy subsets allows us to define fuzziness measures by comparing a fuzzy subset with its complement, with the closest (in some sense) crisp set or with the equilibrium ([6]). However, this restriction to the set of real numbers can lead to the loss of some important information about this difference, namely it does not distinguish whether differences occur in low or high membership degrees.

For the reader's convenience we introduce some basic definitions.

Definition 1. By a fuzzy subset A of the universe Ω we understand a mapping $A : \Omega \to [0; 1]$.

If no confusion can arise, we speak simply of fuzzy sets rather than fuzzy subsets of a given universe. The set of all fuzzy subsets of Ω will be denoted by $F(\Omega)$. We say that a fuzzy set A is a subset of a fuzzy set B if for these functions the inequality $A \leq B$ holds. In the usual way we understand the α-cuts of sets in $F(\Omega)$, i.e. if $\alpha \in (0; 1]$, then the α-cut of a fuzzy set A is the (crisp) set

$$A_\alpha = \{x \in \Omega; A(x) \geq \alpha\}.$$

2000 *Mathematics Subject Classification.* 04A72.
Key words and phrases. Fuzzy set, divergence measure, fuzzy divergence measure.
The research in this paper is supported by Grants DGE-97-PB-1286 (the first and the third author) and VEGA-1/7625/20 (the second author).

If we assume some kind of topology on Ω, then the zero cut of A is the set

$$A_0 = cl\{x \in \Omega; A(x) \geq 0\},$$

where cl is the closure operator.

The intersection and union of fuzzy sets can be defined using an arbitrary triangular norm. Nevertheless, we will work only with the minimum triangular norm here, that means, by the intersection of fuzzy sets A and B we will understand the fuzzy set $A \cap B = \min\{A, B\}$, and by their union the fuzzy set $A \cup B = \max\{A, B\}$.

In several previous papers ([5], [7], [8]) we have introduced and studied a way to quantify the degree of difference between two fuzzy sets by a real function called a divergence measure, which has as its particular cases the usual distances between fuzzy sets already known and used ([4]).

The measure of difference between two fuzzy sets was defined on the basis of the following natural properties:
1) It should be a nonnegative and symmetric function of two fuzzy sets,
2) it should become zero if the two sets coincide,
3) it should decrease if the two sets become more similar in some sense.

While it is easy to formulate analytically the first and the second condition, the third one depends on the formalization of the similarity concept. A possible approach is based on the fact that if a fuzzy set C is added (in the sense of a union) to both A and B, two sets which are closer to each other are obtained; the same should hold for the intersection.

Definition 2. Let Ω be a universe. A mapping $D : F(\Omega)^2 \to R$ is called a divergence measure if for each $A, B, C \in F(\Omega)$ there is:

$$D(A, B) = D(B, A), D(A, A) = 0$$

and

$$\max\{D(A \cup C, B \cup C), D(A \cap C, B \cap C)\} \leq D(A, B).$$

It is also possible that a divergence measure is not defined on the whole set $F(\Omega)^2$, but just on some subset of this product.

The nonnegativity of D follows from the second and the third property, where for the fuzzy set C we put the empty set, i.e. $C(x) = 0$ for each $x \in \Omega$.

A natural candidate for the divergence measure in case of a finite universe (or in case we work only with finite fuzzy sets) is the Hamming distance defined by the formula

$$D(A, B) = \sum_{x \in \Omega} |A(x) - B(x)|.$$

If (Ω, μ) is a measurable space, than another example of a divergence measure which works with integrable fuzzy sets is

$$D(A, B) = \int_{\Omega} |A - B| d\mu.$$

A common disadvantage of this approach is that the divergence between fuzzy sets is expressed by a single real number not accounting on which level the difference between given fuzzy sets is realized.

Example 1. Let $\Omega = \{x; y\}$, let A, B and C be the following fuzzy subsets of Ω:

$$A(x) = 0, A(y) = 1, B(x) = \frac{1}{2}, B(y) = 1, C(x) = 0, C(y) = \frac{1}{2}.$$

Using the Hamming distance we obtain

$$D(A, B) = D(A, C) = \frac{1}{2}.$$

The divergence between A and B in the previous example is the same as the one between A and C. On the other hand, the α-cuts of A and B for $\alpha > \frac{1}{2}$ (i.e. the "important" α-cuts) are the same, which is not true for the fuzzy sets A and C. From this point of view we can require that the divergence between A and B should be smaller than between A and C.

We try to introduce a fuzzy quantity that would reflect the above mentioned difference as well as fulfill properties analogous to those from Definition 2.

Let A be a fuzzy set, let $\alpha \in [0; 1]$. By the symbol A^α we will denote the fuzzy set

$$A^\alpha(x) = \begin{cases} A(x) & \text{if } x \in A_\alpha \\ 0 & \text{otherwise,} \end{cases}$$

where A_α is the α-cut of A. Note that $A = A^0$.

We will also need the definition of a pseudoinverse to a non-increasing function defined on the unit interval.

Definition 3. Let f be a nonnegative non-increasing real function defined on the interval $[0; 1]$. Its pseudoinverse is the function $f^{(-1)} : [0; \infty) \to [0; 1]$ for which $f^{(-1)}(x) = \sup\{r; f(r) > x\}$, with the convention $\sup \emptyset = 0$.

The notion of pseudoinverse can be defined in much more general context (see [10]). For our purpose this definition will be sufficient.

Let now D be an arbitrary divergence measure, let A, B be fuzzy sets such that $D(A^\alpha, B^\alpha)$ exists for each $\alpha \in [0; 1]$. For these sets we can construct the function $\varphi_{A,B}$ in the following way:

$$\varphi_{A,B}(\alpha) = \sup\{D(A^\omega, B^\omega); \omega \geq \alpha\}$$

Obviously the function $\varphi_{A,B}$ depends also on the chosen divergence measure D and this should be reflected in the notation. As we work with a fixed divergence measure, to keep the notation simple we allow this small inaccuracy.

Proposition 1. *For any fuzzy sets A and B the function $\varphi_{A,B}$ is non-increasing.*

Proof. If $\alpha < \beta$, then evidently

$$\{D(A^\omega, B^\omega); \omega \geq \alpha\} \supseteq \{D(A^\omega, B^\omega); \omega \geq \beta\},$$

and the least upper bounds of these sets are therefore in the same order. This yields the required inequality $\varphi_{A,B}(\alpha) \geq \varphi_{A,B}(\beta)$. □

So we can apply the operation of a pseudoinverse to this function.

Definition 4. The fuzzy set $\Delta(A,B) = \varphi_{A,B}^{(-1)}$ will be called the fuzzy divergence measure between A and B,

This function can be considered as a fuzzy quantity $\overline{a}$ corresponding to the linguistic construction "a number not much greater than a". Applying the transformation $\tau(\Delta(A,B)) = 1 - \Delta(A,B)$ we obtain exactly the well-known statistical representation of a positive fuzzy number (see e.g. [3]).

In the following we will show that Δ has similar properties to those of a (crisp) divergence measure from Definition 2, Therefore it can be considered as its generalization. In the following we suppose that all the fuzzy sets we work with admit their mutual fuzzy divergence measure.

Proposition 2. *For all* $A, B \in F(\Omega)$ *there is* $\Delta(A,B) = \Delta(B,A)$.

This statement is a direct consequence of the symmetry from the definition of a divergence measure. Therefore also $\varphi_{A,B} = \varphi_{B,A}$ holds.

The following property expresses the fact that the fuzzy divergence of two sets that coincide is a fuzzy quantity corresponding to the representation of zero.

Proposition 3. *For any* $A \in F(\Omega)$ *there is* $\Delta(A,A)(0) = 1, \Delta(A,A)(x) = 0$ *for all* $x > 0$.

Proof. As $D(A^\omega, A^\omega) = 0$ for all $\omega \in [0;1]$ we have $\varphi_{A,A}(\alpha) = 0$ for all $\alpha \in [0;1]$. Then

$$\Delta(A,A)(0) = \varphi_{A,A}^{(-1)}(0) = \sup\{\alpha \in [0;1]; \varphi_{A,A}(\alpha) \geq 0\} = 1.$$

If $x > 0$, then

$$\{\alpha; \varphi_{A,A}(\alpha) \geq x\} = \emptyset,$$

and hence $\Delta(A,A)(x) = 0$. □

Proposition 4. *For all* $A, B, C \in F(\Omega)$ *there is*

$$\max\{\Delta(A \cap C, B \cap C), \Delta(A \cup C, B \cup C)\} \subseteqq \Delta(A,B).$$

Proof. As the pseudoinverse for a non-increasing function is an order-preserving operation, it is sufficient to show that there is

$$\varphi_{A\cap C, B\cap C} \leq \varphi_{A,B} \quad \text{and} \quad \varphi_{A\cup C, B\cup C} \leq \varphi_{A,B}.$$

If $\omega \in [0;1]$ and $x \in \Omega$, then clearly $(A \cap B)(x) = \min\{A(x), B(x)\} \geq \omega$ if and only if both $A(x) \geq \omega$ and $B(x) \geq \omega$. Hence $(A \cap B)^\omega = A^\omega \cap B^\omega$.

This means that for any $\alpha \in [0;1]$ we have

$$\varphi_{A\cap C, B\cap C}(\alpha) = \sup\{D((A \cap C)^\omega, (B \cap C)^\omega); \omega \geq \alpha\} =$$
$$= \sup\{D(A^\omega \cap C^\omega, B^\omega \cap C^\omega); \omega \geq \alpha\}.$$

As D is a divergence measure, due to its properties the last term is less or equal to

$$\sup\{D(A^\omega, B^\omega); \omega \geq \alpha\} = \varphi_{A,B}(\alpha).$$

Thus the first required inequality is shown. The other can be proved the same way, using the property $(A \cup B)^\omega = A^\omega \cup B^\omega$. □

Now we will return to Example 1 and show that a fuzzy divergence measure provides us with more information comparing to the crisp one.

Example 2. Let $\Omega = \{x; y\}$, let A, B and C have the same meaning as in Example 1. We will find $\Delta(A, B)$ and $\Delta(A, C)$ with Δ based on the Hamming distance.

It is easy to verify that the functions $\varphi_{A,C}, \varphi_{B,C}$ are the following:

$$\varphi_{A,C}(\alpha) = 1 \quad \text{for all} \quad \alpha \in [0; 1],$$

$$\varphi_{B,C}(\alpha) = \begin{cases} \frac{1}{2} & \text{if } \alpha \in [0; \frac{1}{2}] \\ 0 & \text{if } \alpha \in (\frac{1}{2}; 1]. \end{cases}$$

Then using the pseudoinverses of these functions we have

$$\Delta(A, C)(x) = \varphi_{A,C}^{(-1)}(x) = \begin{cases} 1 & \text{if } x \in [0; 1) \\ 0 & \text{if } x \in [1; \infty), \end{cases}$$

$$\Delta(B, C)(x) = \varphi_{B,C}^{(-1)}(x) = \begin{cases} \frac{1}{2} & \text{if } x \in [0; \frac{1}{2}) \\ 0 & \text{if } x \in [\frac{1}{2}; \infty). \end{cases}$$

We see that while the divergence measure based on the Hamming distance is the same (Example 1), their fuzzy divergence measures are different. Moreover, $\Delta(B, C) \leq \Delta(A, C)$, what reflects the fact, that the differences between A and C are on higher membership degree, i.e. in most applications should be considered as more remarkable.

References

[1] Bouchon-Meunier, B., Rifqui, M., Bothorel, S., *Towards general measures of comparisons of objects*, Fuzzy Sets and Systems **84** (1996), 143-153.

[2] Chen, S.M., Yeh, M.S., Hsiao, P.Y., *A comparison of similarity measures of fuzzy values*, Fuzzy Sets and Systems **72** (1995), 79-89.

[3] Klement, E. P., *Integration of fuzzy-valued functions*, Rev. Roumaine Math. Pures Appl. **30** (1985), 375-384.

[4] Klir, G.J., Fogler, T.A., *Fuzzy Sets, Uncertainity and Information*, Prentice Hall International, 1998.

[5] Montes, S., *Partitions and divergence measures in fuzzy models*, Ph.D. Thesis, University of Oviedo (1998).

[6] Montes, S., Couso, I., Gil, P., *Relationships between Uncertainity Measures, Fuzziness Measures and Divergence Measures*, Proceedings of EUSFLAT-ESTYLF Joint Conference, Palma de Mallorca (1999), 227-230.

[7] Montes, S., Gil, P., *Some classes of divergence measures between fuzzy sets and fuzzy partitions*, Mathware and Soft Computing **5** (1999), 253-265.

[8] Montes, S., Gil, P. Bertoluzza, C., *Divergence between fuzzy sets and fuzziness*, Proceedings of IPMU'98 Conference, Paris (1998), 1030-1035.

[9] Pappis, C.P., Karacapilidis, N.I., *A comparative assessment of measures of similarity of fuzzy values*, Fuzzy Sets and Systems **56** (1993), 171-174.

[10] Viceník, P., *A note to a construction of t-norms based on pseudo-inverses of monotone functions*, Fuzzy Sets and Systems **104** (1999), 15-18.

[11] Wang, X., De Baets, B., Kerre, E., *A comparative study of similarity measures*, Fuzzy sets and systems **73** (1995), 259-268.

(Received August 28, 2000)

Dept. of Statistics and O.R.
Faculty of Sciences
Oviedo University
Calvo Sotelo s/n
33071 Oviedo
SPAIN
Dept. of Mathematics
Matej Bel University
Tajovského 40
974 01 Banská Bystrica
SLOVAKIA

E-mail address: janis@fpv.umb.sk
{couso,smr}@pinon.ccu.uniovi.es

Acta Univ. M. Belii
Math. no. 8(2000), pp. 27-33

STRICT ORDER–BETWEENNESSES

Judita Lihová

Abstract. In this note necessary and sufficient conditions for a strick ternary relation to be a strict order–betweenness are given. As an application, a characterization of lattices of convex subsets of posets is obtained.

Introduction

In every poset, we can introduce a ternary relation r called order–betweenness in the following way:

$(a, b, c) \in r$ if and only if $a \leq b \leq c$ or $c \leq b \leq a$.

G. Birkloff proposed (cf. the 2nd edition of his Lattice theory, 1948, Problem 1) to search for axioms for a ternary relation to be an order - betweenness. Such system of axioms was found by M. Altwegg [1], M. Sholander [2] gave an alternative system of axioms.

In certain cases, it is more convenient to handle with strict ternary relations, as sets of triples of different elements. E.g., if we define a ternary relation r' in a poset $\mathbb{A} = (A, \leq)$ to be a strict order - betweenness provided that

$(a, b, c) \in r'$ if and only if $a < b < c$ or $c < b < a$,

the convexity of a subset X of A means that

$(a, b, c) \in r', a, c \in X$ imply $b \in X$.

In this note we give necessary and sufficient conditions for a strict ternary relation to be a strict order–betweenness. As a consequence, a characterization of lattices of convex subsets of posets is obtained. An alternative characterization can be found in [4].

1. Strict order–betweennesses

M. Altwegg proved the following theorem (cf. [1]):

1.1. Theorem. *Let M be a nonempty set, ζ a ternary relation in M. Then there exists a partial order $\leq$ in M with*

$(a, b, c) \in \zeta$ *iff* $a \leq b \leq c$ *or* $a \geq b \geq c$

if and only if ζ satisfies:

(Z_1) $(x, x, x) \in \zeta$ *for each* $x \in M$,

2000 *Mathematics Subject Classification.* 06A06, 06B99.
Key words and phrases. (Strict) order–betweenness, convex subset.
The author was a supported by the Slovak VEGA Grant No. 1/7468/20.

(Z_2) *$(x,y,z)\in\zeta$ implies $(z,y,x)\in\zeta$,*
(Z_3) *$(x,y,z)\in\zeta$ implies $(x,x,y)\in\zeta$,*
(Z_4) *$(x,y,x)\in\zeta$ implies $x=y$,*
(Z_5) *$(x,y,z),(y,z,u)\in\zeta, y\neq z$ imply $(x,y,u)\in\zeta$,*
(Z_6) *if $x_0,x_1,...,x_n,x_{n+1}=x_0,x_{n+2}=x_1$ is a sequence of elements of M such that $(x_{i-1},x_{i-1},x_i)\in\zeta$ and $(x_{i-1},x_i,x_{i+1})\notin\zeta$ for each $i\in\{1,...,n+1\}$ (with a positive integer n), then n is odd.*

Let M be a nonempty set, ξ a ternary relation in M. We will refer to ξ as a strict ternary relation, if it contains only triples of different elements.

Consider the following conditions concerning a strict ternary relation ξ:

(S) $(x,y,z)\in\xi$ implies $(z,y,x)\in\xi$,
(T) $(x,y,z)\in\xi,(y,z,u)\in\xi$ imply $(x,y,u)\in\xi$,
(R) $(x,y,z)\in\xi,(x,z,t)\in\xi$ imply $(y,z,t)\in\xi$,
(X) $(x,y,z)\in\xi,(x',y,z')\in\xi$ imply $(x,y,z')\in\xi$ or $(x,y,x')\in\xi$,
(F) $(x,y,z)\in\xi,(y,u,v)\in\xi$ imply $(x,y,u)\in\xi$ or $(z,y,u)\in\xi$,
(C) $(x,y,z)\in\xi,(y,u,v)\in\xi$ imply $(v,y,x)\in\xi$ or $(z,y,u)\in\xi$,
(I) $(x,y,z)\in\xi$ implies $(x,z,y)\notin\xi$,
(O) if $(a_0,y_1,a_1),(a_1,y_2,a_2),...,(a_{n-1},y_n,a_n)$ is a sequence of elements of ξ such that $(a_{i-1},a_i,a_{i+1})\notin\xi$ for each $i\in\{1,...,n-1\}$ and $(a_0,y_1,a_1)=(a_{n-1},y_n,a_n)$, then n is odd.

1.2. Definition. *By a strict order–betweenness in a set M a strict ternary relation ξ in M satisfying*

$$(a,b,c)\in\xi \text{ iff } a<b<c \text{ or } a>b>c$$

for a partial order $\leq$ in M, will be meant.

We will prove the following theorem.

1.3. Theorem. *Let ξ be a strict ternary relation in a nonempty set M. The following conditions are equivalent:*

(1) *ξ is a strict order–betweenness,*
(2) *ξ satisfies $(S),(T),(X),(F)$ and (O),*
(3) *ξ satisfies $(S),(X),(C)$ and (O).*

First we want to clear the relations between the conditions $(S)-(O)$. The following statements are easy to prove for any strict ternary relation ξ.

1.4. Lemma. *If ξ satisfies (R), then it satisfies (I).*

1.5. Lemma. *If ξ satisfies $(S),(F),(I)$, then it satisfies also (R).*

1.6. Lemma. *If ξ satisfies $(S),(C)$, then it satisfies $(R),(F),(T)$, too.*

1.7. Lemma. *If ξ satisfies $(S),(T),(F)$, then it satisfies also (C).*

1.8. Corollary. *The following assertions are equivalent:*

(a) *ξ satisfies (S) and (C),*
(b) *ξ satisfies $(S),(T)$ and (F),*
(c) *ξ satisfies $(S),(T),(R),(F),(C),(I)$.*

In the following two lemmas we suppose that ξ is a strict ternary relation in a set M satisfying all conditions $(S)-(O)$. Let us define a ternary relation ζ by

$$\zeta = \xi \cup \{(u,u,u) : u \in M\} \cup \{(u,v,w) : u = v \neq w \text{ or } u \neq v = w$$
$$\text{and } u, w \in \{x,y,z\} \text{ for some } (x,y,z) \in \xi\}.$$

The aim is to show that ζ fulfils (Z_6).

1.9. Lemma. *Let $x_0, x_1, ..., x_k \in M$ (k is a prositive integer) and $(x_{i-1}, x_{i-1}, x_i) \in \zeta$ for each $i \in \{1, ..., k\}, (x_{i-1}, x_i, x_{i+1}) \notin \zeta$ for each $i \in \{1, ..., k-1\}, x_0 \neq x_1$. Then there exist $(a_0, y_1, a_1), (a_1, y_2, a_2), ..., (a_{k-1} y_k, a_k) \in \xi$ such that $(a_{i-1}, a_i, a_{i+1}) \notin \xi$ for each $i \in \{1, ..., k-1\}$, further either $x_0 = a_0, x_1 \in \{y_1, a_1\}$ or $x_0 = y_1, x_1 = a_1$ and simultaneously either $x_{k-1} = a_{k-1}, x_k \in \{y_k, a_k\}$ or $x_{k-1} = y_k, x_k = a_k$ holds.*

Proof. We will proceed by induction on k. If $k = 1$, the assertion is evident. Let us remark that we need here the assumption $x_0 \neq x_1$. If $k > 1$, then the relation $x_0 \neq x_1$ is implied by $(x_0, x_1, x_2) \notin \zeta$. Suppose that the assertion is true for a positive integer k. We will prove it for $k+1$. So let $x_0, x_1, ..., x_k, x_{k+1} \in M$ satisfy the above assumptions and let $(a_0, y_1, a_1), ..., (a_{k-1}, y_k, a_k)$ be a sequence corresponding to $x_0, x_1, ..., x_k$ by induction hypothesis. We have the following possibilities:

1) $x_{k-1} = a_{k-1}, x_k = y_k$,
2) $x_{k-1} = a_{k-1}, x_k = a_k$,
3) $x_{k-1} = y_k, x_k = a_k$.

Further, since $(x_k, x_k, x_{k+1}) \in \zeta$ and $x_k \neq x_{k+1}$, because $(x_{k-1}, x_k, x_{k+1}) \notin \zeta$, there exists $t \in M$ with

I) $(x_k, t, x_{k+1}) \in \xi$ or
II) $(t, x_k, x_{k+1}) \in \xi$ or
III) $(t, x_{k+1}, x_k) \in \xi$.

We proceed combining the cases 1) - 3) with I) - III). Let us suppose that 1) and I) occur. The relations $(x_{k-1}, x_k, a_k) = (a_{k-1}, y_k, a_k) \in \xi, (x_k, t, x_{k+1}) \in \xi$ imply $(a_k, x_k, t) \in \xi$ by (C). Then using (T) we obtain $(a_k, x_k, x_{k+1}) \in \xi$. If we show that $(a_{k-1}, a_k, x_{k+1}) \notin \xi$, then adding (a_k, x_k, x_{k+1}) to the sequence $(a_0, y_1, a_1), ..., (a_{k-1}, y_k, a_k)$ we obtain such a sequence as we need. If it were $(a_{k-1}, a_k, x_{k+1}) \in \xi$, we would have $(x_k, a_k, x_{k+1}) = (y_k, a_k, x_{k+1}) \in \xi$ by (R), contrary to the above proved $(a_k, x_k, x_{k+1}) \in \xi$. Combining 1) with II) or III), we can proceed analogously. It is easy to verify that if I) and any of the cases 2), 3) occur, we can add (x_k, t, x_{k+1}), while if III) and simultaneously 2) or 3) occur, we can add (x_k, x_{k+1}, t) to the sequence $(a_0, y_1, a_1), ..., (a_{k-1}, y_k, a_k)$. Now let us suppose that II) and 2) occur. Since $(x_{k-1}, x_k, x_{k+1}) \notin \xi$, we have also $(y_k, x_k, x_{k+1}) \notin \xi$, by (T). Using (C) we obtain $(a_{k-1}, x_k, t) \in \xi$. Now consider the sequence $(a_0, y_1, a_1), ..., (a_{k-2}, y_{k-1}, a_{k-1})$,
$(a_{k-1}, x_k, t), (t, x_k, x_{k+1})$ of elements of ξ. We will show that $(a_{k-2}, a_{k-1}, t) \notin \xi$, $(a_{k-1}, t, x_{k+1}) \notin \xi$. The first relation follows from $(a_{k-2}, a_{k-1}, a_k) \notin \xi$ by (R). If it were $(a_{k-1}, t, x_{k+1}) \in \xi$, we would have $(a_{k-1}, t, x_k) \in \xi$ by (R), contrary to $(a_{k-1}, x_k, t) \in \xi$. Let us notice that in this case we change the last member of the sequence $(a_0, y_1, a_1), ...,$ (a_{k-1}, y_k, a_k) and add a new one. If $k = 2$, we have $(x_0, x_1, t), (t, x_1, x_2)$, as we need. The remaining case, if II) and 3) occur, can be analysed analogously.

1.10 Lemma. The relation ζ fulfils (Z_6).

Proof. Let $x_0, x_1, ..., x_n,\ x_{n+1} = x_0,\ x_{n+2} = x_1$ be a sequence of elements of M such that $(x_{i-1},\ x_{i-1},\ x_i) \in \zeta$ and $(x_{i-1},\ x_i,\ x_{i+1}) \notin \zeta$ for each $i \in \{1, ..., n\}$. The previous lemma ensures the existence of a sequence $(a_0, y_1, a_1), ..., (a_n, y_{n+1}, a_{n+1})$ of elements of ξ such that $(a_{i-1}, a_i, a_{i+1}) \notin \xi$ for each $i \in \{1, ..., n\}$ and one of the following conditions $a), b), c)$ is satisfied and simultancously one of the possibilities $\alpha), \beta), \gamma)$, occurs:

$a)$ $x_0 = a_0,\ x_1 = y_1,$
$b)$ $x_0 = a_0,\ x_1 = a_1,$
$c)$ $x_0 = y_1,\ x_1 = a_1,$

$\alpha)$ $x_n = a_n,\ x_0 = y_{n+1},$
$\beta)$ $x_n = a_n,\ x_0 = a_{n+1},$
$\gamma)$ $x_n = y_{n+1},\ x_0 = a_{n+1}.$

In each of the cases $a)$ and $\beta)$, $a)$ and $\gamma)$, $b)$ and $\beta)$, $b)$ and $\gamma)$, we take the sequence $(a_0, y_1, a_1), ..., (a_n, y_{n+1}, a_{n+1}), (a_{n+1}, y_1, a_1) = (a_0, y_1, a_1)$. In the cases $a)$ and $\alpha)$, $b)$ and $\alpha)$ we take the sequence $(a_0, y_1, a_1), ..., (a_n, y_{n+1}, a_{n+1}), (a_{n+1}, a_0, a_1),$ $(a_1, y_1, a_0), (a_0, y_1, a_1)$. If $c)$ and $\beta)$ or $c)$ and $\gamma)$ occur, we take $(a_0, y_1, a_1), ...,$ (a_{n-1}, y_n, a_n), $(a_n, a_{n+1}, a_0), (a_0, y_1, a_1)$. Finally, if $c)$ and $\alpha)$ occur, we take the sequence $(a_0, y_1, a_1), ..., (a_{n-1}, y_n, a_n), (a_n, y_{n+1}, a_0), (a_0, y_1, a_1)$. Each of these sequences satisfies the assumptions of the condition (O). We will show it, e.g., in the last case. We have $(a_n, y_{n+1}, a_{n+1}) \in \xi, (a_0, y_{n+1}, a_1) = (a_0, y_1, a_1) \in \xi$, so that $(a_n, y_{n+1}, a_0) \in \xi$ by (X), because $(a_n, y_{n+1}, a_1) = (x_n, x_0, x_1) = (x_n, x_{n+1}, x_{n+2}) \notin \xi$. Further we will show $(a_{n-1}, a_n, a_0) \notin \xi, (a_n, a_0, a_1) \notin \xi$. If it were $(a_n, a_0, a_1) \in \xi$, we would have $(y_{n+1}, a_0, a_1) \in \xi$ by (R), which contradicts $(a_0, y_{n+1}, a_1) = (a_0, y_1, a_1) \in \xi$. Let us suppose that $(a_{n-1}, a_n, a_0) \in \xi$. Using $(a_n, y_{n+1}, a_{n+1}) \in \xi$ we obtain (a_0, a_n, y_{n+1}) by (C), because $(a_{n+1}, a_n, a_{n-1}) \notin \xi$. But this is a contradiction, as we have proved $(a_n, y_{n+1}, a_0) \in \xi$. Now using (O) we conclude that n is odd.

Proof of theorem 1.3. It is easy to see that (1) implies (2). Further, (2) implies (3) by 1.8. Now let ξ satisfy $(S), (X), (C)$ and (O). Then ξ satisfies all conditions $(S) - (O)$, again by 1.8. Let ζ be defined as before 1.9. It is easy to verify that ζ satisfies $(Z_1) - (Z_5)$. By lemma 1.10 it satisfies (Z_6), too. Theorem 1.1 ensures the existence of a partial order $\leq$ in M such that $(a, b, c) \in \zeta$ if and only if either $a \leq b \leq c$ or $a \geq b \geq c$ holds. Obviously $(a, b, c) \in \xi$ is equivalent to $a < b < c$ or $a > b > c$, so that ξ is a strict order–betweenness.

The following examples show that the system of conditions given in (2) of 1.3 and (3) of 1.3, respectively, is independent. In each of these examples we point out, which of the conditions $(S) - (O)$ are not satisfied.

1.11 Example. Let $M = \{a, b, c\}$, $\xi = \{(a, b, c)\}$. Then ξ doesn't satisfy (S).

1.12 Example. Let $M = \{a, b, c\}$, $\xi = \{(a, b, c), (a, c, b), (c, b, a), (b, c, a)\}$. Then ξ doesn't satisfy $(T), (R), (C), (I)$.

1.13 Example. Let $M = \{a, b, c, d, e\}$, $\xi = \{(b, c, d), (a, c, e), (d, c, b), (e, c, a)\}$. Then ξ doesn't satisfy (X).

1.14 Example. Let $M = \{a,b,c,d,e\}$, $\xi = \{(a,b,c),(b,d,e),(c,b,a),(e,d,b)\}$. Then ξ doesn't satisfy (F) and (C).

1.15 Example. Let $M = \{a,b,c,d,e,f\}$, $\xi = \{(a,b,c),(c,d,e),(e,f,a),(c,b,a),(e,d,c),(a,f,e)\}$. Then ξ doesn't satisfy (O).

2. Characterization of Conv $\mathbb{A}$

Theorem 1.3 enables us to give a characterization of lattices of convex subsets of partially ordered sets. For a partially ordered set $\mathbb{A} = (A, \leq)$ let Conv $\mathbb{A}$ denote the system of all convex subsets of A. It is easy to see that (Conv $\mathbb{A}, \subseteq)$ is a complete atomistic lattice (atomisticity means that every element is a join of atoms).

2.1 Theorem. *Let $\mathbb{L} = (L, \wedge, \vee, \leq)$ be a complete atomistic lattice, card $L > 1$. Further, let M be the set of all atoms of $\mathbb{L}$, ξ the ternary relation in M defined by*

$$(a,b,c) \in \xi \Longleftrightarrow b < a \vee c,\ b \neq a,\ b \neq c.$$

The following conditions are equivalent:

(I) *$\mathbb{L}$ is isomorphic to Conv $\mathbb{A}$ for a partially ordered set $\mathbb{A}$;*

(II) *ξ satisfies $(T),(X),(F),(O)$ and*

(K) *$a \leq \sup X,\ X \subseteq M,\ a \in M - X$ imply $(x_1, a, x_2) \in \xi$ for some $x_1, x_2 \in X$;*

(III) *ξ satisfies $(X),(C),(O)$ and (K).*

Proof. Since the relation ξ is evidently symmetric, the conditions $(II),(III)$ are equivalent by 1.8. To prove $(I) \Rightarrow (II)$, let φ be an isomorphism of $\mathbb{L}$ onto Conv $\mathbb{A}$ for a partially ordered set $\mathbb{A} = (A, \leq^*)$. As atoms of the lattice Conv $\mathbb{A}$ are just the one-element subsets of A, the mapping $\varphi' : M \to A$ defined by

$$\varphi'(x) = a \Longleftrightarrow \varphi(x) = \{a\}$$

is a bijection of M onto A. Evidently $(x,y,z) \in \xi$ means that either $\varphi'(x) <^* \varphi'(y) <^* \varphi'(z)$ or $\varphi'(z) <^* \varphi'(y) <^* \varphi'(x)$ holds. Consider the partial order $\leq'$ in M defined in such a way that φ' is an isomorphism of $(M, \leq')$ onto $\mathbb{A}$. Then we have $(x,y,z) \in \xi$ if and only if $x <' y <' z$ or $z <' y <' x$ holds, so that ξ is a strict order–betweeenness. Using theorem 1.3 we obtain that ξ satisfics $(T),(X),(F)$ and (O). It remains to show that (K) is satisfied. So let $a \leq \sup X,\ X \subseteq M,\ a \in M - X$. Then $\varphi'(a)$ belongs to the convex hull of $\{\varphi'(x) : x \in X\}$ in $\mathbb{A}$. Consequently there exist $x_1, x_2 \in X$ with $\varphi'(x_1) \leq^* \varphi'(a) \leq^* \varphi'(x_2)$. Since $a \notin X$, we have $x_1 <' a <' x_2$ and hence $(x_1, a, x_2) \in \xi$. We are going to prove $(II) - (I)$. So let ξ satisfy $(T),(X),(F),(O)$ and (K). Since ξ is also symmetric, we have $(a,b,c) \in \xi$ if and only if $a <^* b <^* c$ or $c <^* b <^* a$ for a partial order $\leq^*$ in M, by 1.3. We will show that $\mathbb{L}$ is isomorphic to Conv $(M, \leq^*)$. Notice that a subset X of M is convex if and only if $(x_1, a, x_2) \in \xi,\ x_1, x_2 \in X$ imply $a \in X$. If $a \in L$, let M_a denote the set $\{p \in M : p \leq a\}$. To verify that the set M_a is convex in $(M, \leq^*)$, let $(u,x,v) \in \xi,\ u,v \in M_a$. But then $x < u \vee v \leq a$, hence $x \in M_a$. Obviously $a \leq b$

implies $M_a \subseteq M_b$. Since $a = \sup M_a$, the converse implication holds, too. Finally, let X be any subset of M, convex in $(M, \leq^*)$ and let $a = \sup X$. We will prove $M_a = X$. The inclusion $X \subseteq M_a$ is evident. Let us suppose that there exists an element $p \in M_a - X$. The condition (K) ensures the existence of $x_1, x_2 \in X$ with $(x_1, p, x_2) \in \xi$. In view of the fact that X is convex we have $p \in X$, a contradiction. The proof is complete.

To show that no of the conditions given in (II) and (III), respectively, can be omitted, consider the following examples.

2.2 Example. Let $\mathbb{L}$ be as in Fig. 1. Then evidently $\xi = \{(a,b,c),(c,b,a)\}$ and it satisfies all conditions $(S)-(O)$, but it doesn't satisfy (K). Namely $a \leq \sup\{b,c,d\}$, while $(b,a,c),(b,a,d),(c,a,d) \notin \xi$.

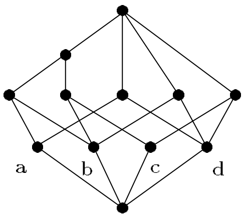

Fig. 1

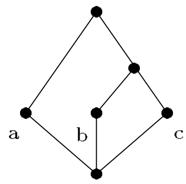

Fig. 2

2.3 Example. Let $\mathbb{L}$ be as in Fig. 2. Then evidently ξ is that of example 1.12. Hence it satisfies $(X),(F),(O)$ and also (K), while $(T),(R),(C)$ and (I) are not satisfied.

2.4 Example. Let $\mathcal{L}$ be the system of all subsets X of the set $\{a,b,c,d,e\}$ satisfying

$$b,d \in X \text{ or } a,e \in X \Rightarrow c \in X.$$

Then $(\mathcal{L}, \subseteq)$ is an atomistic lattice and the relation ξ corresponds to that of example 1.13. So it satisfies (K) and all $(S)-(O)$, besides (X).

2.5 Example. Let $\mathcal{L}$ be the system of all subsets X of $\{a,b,c,u,v\}$ with

$$a,c \in X \Rightarrow b \in X,$$

$$v \in X \text{ and } (b \in X \text{ or } c \in X) \Rightarrow u \in X.$$

Then $(\mathcal{L}, \subseteq)$ is an atomistic lattice, $\xi = \{(\{a\},\{b\},\{c\}),(\{c\},\{b\},\{a\}),(\{b\},\{u\},\{v\}),(\{v\},\{u\},\{b\}),(\{c\},\{u\},\{v\}),(\{v\},\{u\},\{c\})\}$. It can be seen easily that from among the conditions $(S)-(O)$ and (K), just (F) and (C) are not satisfied.

2.6 Example. Let $\mathcal{L}$ be the system of all subsets X of the set $\{a,b,c,d,e,f\}$ satisfying

$$a,c \in X \Rightarrow b \in X,$$

$$c,e \in X \Rightarrow d \in X,$$

$$e,a \in X \Rightarrow f \in X.$$

Then $(\mathcal{L}, \subseteq)$ is an atomistic lattice. The relation ξ corresponds to that of example 1.15, hence it satisfies all conditions $(S)-(I)$ and it doesn't satisfy (O). Evidently (K) holds, too.

Another characterization of lattices of convex subsets of partially ordered sets is given in [4]. We refer to such lattices as c-lattices there. It is also proved that each c-lattice is a direct product of directly irreducible c-lattices and directly irreducible c-lattices are described. The construction of all partially ordered sets $\mathbb{B}$ with Conv $\mathbb{B}$ isomorphic to Conv $\mathbb{A}$ for any given partially ordered set $\mathbb{A}$ can be found in [3].

References

[1] M. Altwegg, *Zur Axiomatik der teilweise geordneten Mengen*, Comment. Math. Helv. **24** (1950), 149–155.

[2] M. Sholander, *Trees, lattices, order, and betweenness*, Proceedings of AMS **3** (1952), 369–381.

[3] J. Lihová, *On convexly isomorphic posets*, Czechoslov. Math. Journal, **49(124)** (1999), 135–148.

[4] J. Lihová, *Characterization of lattices of convex subsets of posets*, Czechoslov. Math. Journal **50(125)** (2000), 113–119.

(Received October 2, 2000)

Dept. of Geometry and Algebra
P. J. Šafárik University
Jesenná 5
04154 Košice
SLOVAKIA

E-mail address: lihova@duro.science.upjs.sk

Acta Univ. M. Belii
Math. no. 8(2000), pp. 35-41

EXTREMUM CONDITIONS FOR A DEGENERATED CRITICAL POINT

Peter Maličký

Abstract. For a degenerated critical point of a function of two variables are given necessary and sufficient conditions for a local extremum.

It is well known that a type of a critical point (point of minimum, maximum or saddle point) for a function f of two variables may be determined using partial derivatives of the second order whenever the hessian is nonzero. The case of the zero hessian is considered as complicated and using of derivatives of higher order is recommended. The present paper shows how to use the derivatives of higher order. It was motivated by paper [1] which contains some necessary condition.

For the simplicity we shall assume that the origin $(0,0)$ is a critical point of the function f and we write

$$a_{ij} = \frac{\partial^{i+j} f}{\partial x^i \partial y^j}(0,0).$$

We have

$$a_{10} = a_{01} = 0. \tag{1}$$

We assume the zero hessian

$$a_{20}a_{02} - a_{11}^2 = 0. \tag{2}$$

We first consider the case, when one of the derivatives of the second order is nonzero. So, we have

$$a_{20} \neq 0 \text{ or } a_{02} \neq 0.$$

For the simplicity we assume

$$a_{20} > 0 \tag{3}$$

and we are interested in conditions under which the origin $(0,0)$ is a point of local minimum of f (maximum is impossible). Let the function f has continuous partial derivatives to the fourth order in a neighbourhood of $(0,0)$. Put

$$g(x,y) = f(x + p_1 y + \frac{p_2}{2} y^2, y)$$

2000 *Mathematics Subject Classification.* 26B05, 26B10.
Key words and phrases. Critical point, hessian.

Functions f and g have a local minimum at the point $(0,0)$ simultaneously, because the maps

$$\Phi : (x, y) \mapsto (x + p_1 y + \frac{p_2}{2} y^2, y)$$

and

$$\Psi : (x, y) \mapsto (x - p_1 y - \frac{p_2}{2} y^2, y)$$

are mutually inverse homeomorphisms which preserve the origin. We denote

$$\frac{\partial^{i+j} g}{\partial x^i \partial y^j}(0,0) = b_{ij} \text{ for } i + j \leq 4.$$

Then we have

$$(4) \qquad b_{10} = a_{10} = 0$$
$$(5) \qquad b_{20} = a_{20} > 0$$
$$(6) \qquad b_{01} = a_{01} + a_{10}p_1 = 0$$
$$(7) \qquad b_{02} = a_{02} + 2a_{11}p_1 + a_{20}p_1^2 + a_{10}p_2 = a_{02} + 2a_{11}p_1 + a_{20}p_1^2$$
$$(8) \qquad b_{03} = a_{03} + 3a_{12}p_1 + 3a_{21}p_1^2 + a_{30}p_1^3 + 3a_{11}p_2 + 3a_{20}p_1p_2$$
$$(9) \qquad b_{04} = a_{04} + 4a_{13}p_1 + 6a_{22}p_1^2 + 4a_{31}p_1^3 + a_{40}p_1^4 + 6a_{12}p_2 + \\ + 12a_{21}p_1p_2 + 6a_{30}p_1^2p_2 + 3a_{20}p_2^2$$
$$(10) \qquad b_{11} = a_{11} + a_{20}p_1$$
$$(11) \qquad b_{12} = a_{12} + 2a_{21}p_1 + a_{30}p_1^2 + a_{20}p_2$$

Choose p_1 and p_2 such that

$$(12) \qquad b_{11} = 0$$
$$(13) \qquad b_{12} = 0$$

It means

$$(14) \qquad p_1 = -\frac{a_{11}}{a_{20}}$$
$$(15) \qquad p_2 = \frac{-1}{a_{20}}(a_{12} + 2a_{21}p_1 + a_{30}p_1^2)$$

Put $H_3 = b_{03}$ and $H_4 = b_{04}$, where p_1 and p_2 are defined by (14) and (15). Then (8), (9), (10), (11), (14) and (15) imply

$$(16) \qquad H_3 = a_{03} + 3a_{12}p_1 + 3a_{21} + a_{30}p_1^3$$
$$(17) \qquad H_4 = a_{04} + 4a_{13}p_1 + 6a_{22}p_1^2 + 4a_{31}p_1^3 + a_{40}p_1^4 - 3a_{20}p_2^2$$

By (7), (14) and (2) we have

$$b_{02} = a_{02} - 2a_{11}\frac{a_{11}}{a_{20}} + a_{20}\frac{a_{11}^2}{a_{20}^2} = a_{02} - 2\frac{a_{02}a_{20}}{a_{20}} + a_{20}\frac{a_{02}a_{20}}{a_{20}^2} = a_{02} - 2a_{02} + a_{02} = 0\ .$$

Theorem 1. *Let f be a function of two variables which has continuous partial derivatives to the fourth order in some neighbourhood of the origin which is a degenerated critical point of f. Let*

$$a_{20} > 0 \; .$$

Conditions $H_3 = 0$ and $H_4 \geq 0$ (resp. $H_3 = 0$ and $H_4 > 0$) are necessary (resp. sufficient) for a local minimum of f at the origin.

Proof. Let f has a local minimum at the point $(0, 0)$. Define

$$\varphi(y) = g(0, y)$$

or equivalently

$$\varphi(y) = f(p_1 y + \frac{p_2}{2} y^2, y) \; .$$

Then φ has a local minimum at 0. We have

$$\varphi^{(k)}(0) = \frac{\partial^k g}{\partial y^k}(0,0) = b_{0k} \text{ for } k = 1, 2, 3, 4 \; .$$

By (6) and (16)

$$\varphi'(0) = \varphi''(0) = 0 \; .$$

Therefore conditions

$$H_3 = b_{03} = \varphi'''(0) = 0$$

and

$$H_4 = b_{04} = \varphi^{(4)}(0) \geq 0$$

are necessary. Now, we prove sufficiency. By Taylor's formula we have

$$g(x, y) = g(0, 0) + \sum_{1 \leq i+j \leq 4} \frac{b_{ij}}{i!j!} x^i y^j + r_4(x, y), \text{ where}$$

$$\lim_{\substack{x \to 0 \\ y \to 0}} \frac{r_4(x, y)}{(x^2 + y^2)^2} = 0$$

which implies

$$\lim_{\substack{x \to 0 \\ y \to 0}} \frac{r_4(x, y)}{x^2 + y^4} = 0,$$

The sum $\sum_{1 \leq i+j \leq 4} \frac{b_{ij}}{i!j!} x^i y^j$ does not contain the terms x, y, y^2, xy, y^3 and xy^2 , because the corresponding b_{ij} are zero. The inequality

$$|xy^3| = |y| \sqrt{x^2 y^4} \leq |y| \frac{x^2 + y^4}{2} \leq |y|(x^2 + y^4)$$

shows that the term xy^3 is negligible with respect to (x^2+y^4). Since $x^2 \le x^2+y^4$, terms x^3, x^2y, x^2y^2, x^3y and x^4 are also negligible with respect to (x^2+y^4). Therefore,

$$g(x,y) = g(0,0) + \frac{b_{20}}{2}x^2 + \frac{b_{04}}{24}y^4 + s(x,y),$$

where

$$\lim_{\substack{x\to 0\\ y\to 0}} \frac{s(x,y)}{x^2+y^4} = 0.$$

It proves sufficiency.

Example 1. Put

$$\begin{aligned}
f_1(x,y) &= x^2+4xy+4y^2+6xy^2+9y^4\\
f_2(x,y) &= x^2+4xy+4y^2+6xy^2+12y^3-6xy^3-2y^4\\
f_3(x,y) &= x^2+4xy+4y^2+6xy^2+12y^3+8y^4
\end{aligned}$$

In all cases $p_1=-2,\ p_2=-6$. We have $H_3=-72$ for f_1 , $H_3=0$ and $H_4=24$ for f_2 and $H_3=0$ and $H_4=-24$ for f_3 . So, only the function f_2 has a local minimum at $(0,0)$.

Of course, it may happen that $H_3=H_4=0$.

Example 2. Put

$$\begin{aligned}
f_4(x,y) &= x^2+4xy+4y^2+12xy^2+24y^3+2xy^3+40y^4+13y^5+y^6\\
f_5(x,y) &= x^2+4xy+4y^2+12xy^2+24y^3+2xy^3+40y^4+12y^5+2y^6\\
f_6(x,y) &= x^2+4xy+4y^2+12xy^2+24y^3+2xy^3+40y^4+12y^5
\end{aligned}$$

Then $p_1=-2,\ p_2=-12$ and $H_3=H_4=0$ in all cases. Put

$$g_i(x,y) = f_i(x-2y-6y^2, y) \text{ for } i=4,5,6\ .$$

Then

$$\begin{aligned}
g_4(x,y) &= x^2+2xy^3+y^5+y^6 = (x+y^3)^2+y^5\\
g_5(x,y) &= x^2+2xy^3+2y^6 = (x+y^3)^2+y^6\\
g_6(x,y) &= x^2+2xy^3 = (x+y^3)^2-y^6
\end{aligned}$$

Now, define

$$h_i(x,y) = g_i(x-y^3, y) \text{ for } i=4,5,6\ .$$

Then

$$\begin{aligned}
h_4(x,y) &= x^2+y^5\\
h_5(x,y) &= x^2+y^6\\
h_6(x,y) &= x^2-y^6
\end{aligned}$$

So, only f_5 has a local minimum at the origin.

If $H_3 = H_4 = 0$, then the previous example indicates that a type of a critical point may be determined by the function h defined by

$$h(x,y) = g(x + \frac{p_3}{6}y^3, y) = f(x + p_1 y + \frac{p_2}{2}y^2 + \frac{p_3}{6}y^3, y).$$

In fact, in this case it is possible to define characteristics H_5 and H_6 (if the function f is six times continuously differentiable in some neighbourhood of the origin) and an analog of Theorem 1 in terms of H_5 and H_6 may be proved. However, we omit the details, because H_5 and H_6 contain 16 and 23 terms respectively.

Example 3. Put

$$f_7(x,y) = \begin{cases} x^2 + e^{-\frac{1}{y^2}} \text{ for } y \neq 0 \\ x^2 \text{ for } y = 0 \end{cases}$$

$$f_8(x,y) = \begin{cases} x^2 + e^{-\frac{1}{y^2}} \text{ for } y > 0 \\ x^2 - e^{-\frac{1}{y^2}} \text{ for } y < 0 \\ x^2 \text{ for } y = 0 \end{cases}$$

Functions f_7 and f_8 have the same partial derivatives (of all orders) at the origin, but only f_7 has a local minimum at the origin. It shows that values of partial derivatives of all orders at a critical point need not determine its type.

Now, assume that

$$a_{02} = a_{20} = 0 = a_{02}a_{20} - a_{11}^2$$

Then also

$$a_{11} = 0.$$

Put

$$P_3(x,y) = a_{30}x^3 + 3a_{21}x^2y + 3a_{12}xy^2 + a_{03}y^3 .$$

and

$$P_4(x,y) = a_{40}x^4 + 4a_{31}x^3y + 6a_{22}x^2y^2 + 4a_{13}xy^3 + a_{04}y^4 .$$

Theorem 2. *Let f be a function of two variables which has continuous partial derivatives to the fourth order in some neighbourhood of the origin which is a critical point of f. Let*

$$a_{02} = a_{20} = a_{11} = 0 .$$

Conditions

$$a_{30} = a_{21} = a_{12} = a_{03} = 0$$

and

$$P_4(x,y) \geq 0 \ \ (\textit{resp. } P_4(x,y) > 0 \ \ \textit{whenever } x^2 + y^2 \neq 0\ )$$

are necessary (resp. sufficient) for a local minimum of f at the origin.

Proof. Let the origin be a point of local minimum of f. For arbitrary reals α and β define

$$\varphi(t) = f(\alpha t, \beta t) \ .$$

Then

$$\varphi'(0) = \varphi''(0) = 0 \ .$$

Since the function φ has a local minimum at 0, we have

$$\varphi'''(0) = a_{30}\alpha^3 + 3a_{21}\alpha^2\beta + 3a_{12}\alpha\beta^2 + a_{03}\beta^3 = P_3(\alpha, \beta) = 0$$

and

$$\varphi^{(4)}(0) = P_4(\alpha, \beta) \geq 0 \ .$$

So, the polynomial P_3 is identically zero. Therefore, all its partial derivatives (of all orders) are identically zero. Particularly, all coefficients of P_3 are zero. It proves necessity. Now, we prove sufficiency. By Taylor's formula we have

$$f(x, y) = f(0, 0) + \frac{1}{4!}P_4(x, y) + r_4(x, y) \ ,$$

where

$$\lim_{\substack{x\to 0\\ y\to 0}} \frac{r_4(x, y)}{(x^2 + y^2)^2} = 0.$$

Put

$$c = \min_{x^2+y^2=1} P_4(x, y).$$

Clearly,

$$c > 0$$

and

$$P_4(x, y) \geq c(x^2 + y^2)^2 \text{ whenever } x^2 + y^2 = 1.$$

Since $P_4(x, y)$ and $(x^2+y^2)^2$ are homogeneous polynomials of the same degree, the last inequality holds for all x and y. It means

$$\lim_{\substack{x\to 0\\ y\to 0}} \frac{r_4(x, y)}{P_4(x, y)} = 0.$$

Therefore, $r_4(x, y)$ is negligible with respect to $P_4(x, y)$ and f has a local minimum at the origin.

References

1. Novotný, P., *A note on local extremes*, Studies of University of transport and communications Žilina, Math.-Phys. Series **10** (1995), 41 — 44.

(Received October 12, 2000)

Dept. of Mathematics
Matej Bel University
Tajovského 40
974 01 Banská Bystrica
SLOVAKIA

E-mail address: malicky@fpv.umb.sk

Acta Univ. M. Belii
Math. no. 8(2000), pp. 43-50

GENERIC CHAOS IN METRIC SPACES

Elena Murinová

Abstract. A dynamical system given by a continuous map f from a metric space X into itself is called generically ε-chaotic if the set of Li-Yorke pairs, i.e., the set of points $[x,y] \in X^2$ for which $\liminf_{n\to\infty} \varrho(f^n x, f^n y) = 0$ and simultaneously $\limsup_{n\to\infty} \varrho(f^n x, f^n y) > \varepsilon$ is residual in X^2. If $\varepsilon = 0$, f is called generically chaotic. It is shown that the characterization of generically ε-chaotic maps given by Ľ. Snoha in the interval case can be extended to a large class of metric spaces. While on the interval generic chaos implies generic ε-chaos for some $\varepsilon > 0$, in the paper an example of a convex continuum in the plane is given on which generic chaos does not imply generic ε-chaos for any $\varepsilon > 0$.

1. Introduction.

We will study a dynamical system $(X; f)$ given by a metric space (X, ϱ) and a continuous map $f : X \to X$ (in written $f \in C(X)$). Usually when studying chaoticity of such systems the authors assume that X is compact. Instead, we will only assume that X is complete (even less, see below).

The notion of *chaos* in connection with a map was first used by Li and Yorke [LY] without giving any formal definition. Since then many definitions of chaos appeared, most of them being surveyed in [KS]. Each of them reflects some aspects of the dynamics of those systems which are generally considered to be really 'chaotic'.

The notion of *generic chaos* was introduced by A. Lasota (see [P]). A system $(X; f)$ is generically chaotic if the set of so called Li-Yorke pairs of points, i.e., the set of points $[x,y] \in X^2$ for which $\liminf_{n\to\infty} \varrho(f^n x, f^n y) = 0$ and $\limsup_{n\to\infty} \varrho(f^n x, f^n y) > 0$ is residual in X^2 (i.e., its complement is a first category set in X^2).

J. Piórek [P] in 1985 found examples of generically chaotic interval maps, so it became clear that maps satisfying such a strong definition of chaoticity exist.

Ľ. Snoha [S1] in 1990 gave a full characterization of generically chaotic self-maps of a real compact interval I in terms of behaviour of subintervals of I as well as in terms of topological transitivity. He also introduced the notion of *dense chaos* by requiring that the set of Li-Yorke pairs be dense instead of residual. In [S2] he found a full characterization of densely chaotic interval maps and proved that in the class of piecewise monotone maps with finite number of pieces of monotonicity

2000 *Mathematics Subject Classification.* 37B05, 54H20.

Key words and phrases. Generic chaos, dense chaos, Li-Yorke pair, Baire space, complete metric space, triangular map.

The author has been supported by the Slovak Grant Agency, grant number 1/7424/20.

the notion of generic chaos and that of dense chaos coincide. Finally, in [S3] he generalized some results from [S1] to what he called two-parameter chaos.

The inspiration for the present paper was a concluding remark of Ľ. Snoha from [S2] saying that "Some results concerning the generic chaos can be carried over to the case of continuous self-maps of the compact metric spaces. For example, if for every two balls B_1 and B_2, $\liminf_{n\to\infty} \operatorname{dist}(f^n(B_1), f^n(B_2)) = 0$ and if there is an $a > 0$ such that for every ball B, $\limsup_{n\to\infty} \operatorname{diam} f^n(B) > a$, then f is generically chaotic." The main aim of the present paper is to develop this idea of Ľ. Snoha and to check to what extent his characterization of generically chaotic maps from [S1][1] and that of generically (α, β)-chaotic maps from [S3] can be carried over from the interval to metric spaces.

Before going further we need to discuss the question which metric spaces will be appropriate for us to work with.

First of all, note that the definition of generic chaos has a good sense only if the space X^2 is of second category (i.e., not of first category) in itself because only then a residual set in the space X^2 can reasonably be considered to form a 'majority' of it (usually, in spaces of first category the residuality is not being defined at all). Still, a residual set in a space of second category need not be dense in the space (e.g., take the space $[0,1] \cup (\mathbb{Q} \cap [2,3])$ with the metric inherited from the real line and the set $[0,1]$). But in the definition of generic chaos the residual set of Li-Yorke pairs should be required to be automatically dense, we believe.

Therefore we will require that X^2 be a Baire space — then X^2 is of second category in itself and any residual set in X^2 is automatically dense. (Recall that a space Y is Baire if every open set in Y is of second category in Y or, equivalently, in itself. This is equivalent with the property that the intersection of any countable collection of open dense sets is dense in Y. Another equivalent definition is that any residual set in Y is dense in Y. See, e.g., [HMcC]).)

Of course, it could seem more reasonable to assume something on the space X itself rather than on X^2. First, we should realize that a necessary condition for X^2 to be Baire is that X be Baire. Unfortunately, this is not a sufficient condition — the square of a metric Baire space need not be Baire (see [Kr] or [HMcC]).

The question therefore is what assumptions on X ensure that X^2 be Baire. Here we wish to mention at least that, among others, any one of the following three conditions is sufficient for X^2 to be Baire (see [HMcC, Theorem 2.4, Proposition 1.23, Theorem 5.1]):

(A1) X is a complete metric space.
(A2) X is a G_δ set in a complete metric space.
(A3) X is Baire and separable metric space.

We thus finish our discussion about the assumptions on X: we will assume that *X is a metric space whose square is Baire.* In particular, it is sufficient to assume that X satisfies any one of the above three conditions.

Now let us go to our results. But first recall some definitions. Consider a

[1]He brought the attention of the author of the present paper to a misprint in [S1, Theorem 1.2] — the condition "(h-1) f has a unique ..." should read as "(h-1) f is not constant in any subinterval of I and has a unique ...".

dynamical system $(X; f)$ and $\varepsilon > 0$ and denote

$$C(f) = \left\{ [x,y] \in X^2 : \liminf_{n\to\infty} \varrho(f^n x, f^n y) = 0 \text{ and } \limsup_{n\to\infty} \varrho(f^n x, f^n y) > 0 \right\},$$
$$C(f,\varepsilon) = \left\{ [x,y] \in X^2 : \liminf_{n\to\infty} \varrho(f^n x, f^n y) = 0 \text{ and } \limsup_{n\to\infty} \varrho(f^n x, f^n y) > \varepsilon \right\}.$$

We say that f is *generically* or *densely chaotic* if the set $C(f)$ is residual or dense in X^2, respectively. Similarly, f is *generically* or *densely ε-chaotic* if the set $C(f,\varepsilon)$ is residual or dense in X^2, respectively.

In [S1] it is among others proved that if $f \in C(I)$ where I is a real compact interval then the following are equivalent:

(a) f is generically chaotic,
(b) for some $\varepsilon > 0$, f is generically ε-chaotic,
(c) for some $\varepsilon > 0$, f is densely ε-chaotic,
(d) the following two conditions are fulfilled simultaneously:
  (d1) for every two intervals J_1, J_2, $\liminf_{n\to\infty} \varrho(f^n(J_1), f^n(J_2)) = 0$,
  (d2) there is $\varepsilon > 0$ such that for every interval J, $\limsup_{n\to\infty} \operatorname{diam} f^n(J) > \varepsilon$.

(Moreover, the equivalences $(b) \Leftrightarrow (c) \Leftrightarrow (d)$ hold with the same ε. Further, any generically chaotic function is densely chaotic but not conversely.)

We show that this result can be extended to metric spaces, though not completely (the implication $(c) \Rightarrow (a)$ in the next theorem does not hold with the same ε, contrary to the interval case).

Theorem A. *Let (X,ϱ) be a metric space whose square X^2 is a Baire space and let $f \in C(X)$. Then the following three conditions are equivalent:*

(a) *for some $\varepsilon > 0$, f is generically ε-chaotic,*
(b) *for some $\varepsilon > 0$, f is densely ε-chaotic,*
(c) *the following two conditions are fulfilled simultaneously:*
  (c1) *for every two balls B_1, B_2, $\liminf_{n\to\infty} \varrho(f^n(B_1), f^n(B_2)) = 0$,*
  (c2) *there exists some $\varepsilon > 0$ such that for every ball B, $\limsup_{n\to\infty} \operatorname{diam} f^n(B) > \varepsilon$.*

Moreover, the implications $(a) \Rightarrow (b) \Rightarrow (c)$ hold with the same ε. The implication $(c) \Rightarrow (a)$ does not hold with the same ε, in general. Nevertheless, one can claim that the condition (c) implies that f is generically ε^*-chaotic for any $\varepsilon^* < \varepsilon/2$.

We also show that, contrary to the interval case, in metric spaces generic chaos does not imply generic ε-chaos. Recall that a metric space is called a continuum if it is compact and connected.

Theorem B. *There is a continuum X in the euclidean plane and a map $f \in C(X)$ such that f is generically chaotic but is not generically ε-chaotic for any $\varepsilon > 0$. The continuum X can even be taken to be convex.*

Acknowledgements. The author thank Ľ. Snoha for many stimulating and very helpful discussions as well as for his enormous help with preparation of this paper.

2. Proof of Theorem A and a generalization.

Being inspired by [S3] we are going to prove a result which is more general than Theorem A.

For a dynamical system $(X;f)$ and real numbers α, β define the following subsets of the square X^2:

$$\begin{aligned} C_1(f,\alpha) &= \left\{[x,y]\in X^2 : \liminf_{n\to\infty} \varrho(f^n x, f^n y)\leqslant \alpha\right\}, \\ C_2(f,\beta) &= \left\{[x,y]\in X^2 : \limsup_{n\to\infty} \varrho(f^n x, f^n y) > \beta\right\}, \\ C(f,\alpha,\beta) &= C_1(f,\alpha)\cap C_2(f,\beta)\ . \end{aligned}$$

Since we speak on chaos, it would be reasonable to consider only $0\leqslant\alpha\leqslant\beta<\operatorname{diam}X$ (in particular, $C_1(f,\alpha)=\emptyset$ for $\alpha<0$ and $C_2(f,\beta)=\emptyset$ for $\beta\geqslant\operatorname{diam}X$). Nevertheless, the results will work for any α, β and therefore we will not assume any restrictions on them.

According to [S3] a map $f\in C(X)$ is called *generically* or *densely* (α,β)-*chaotic* if the set $C(f,\alpha,\beta)$ is residual or dense in X^2, respectively.

If $\alpha=0$ or $\alpha=\beta=0$ we sometimes omit them. More precisely, instead of generic or dense $(0,\varepsilon)$-chaos we also shortly speak on generic or dense ε-chaos, respectively and instead of generic or dense $(0,0)$-chaos we simply speak on generic or dense chaos, respectively. Thus, this terminology is in accordance with the fact that for above defined sets $C(f)$ and $C(f,\varepsilon)$ we have $C(f)=C(f,0,0)$ and $C(f,\varepsilon)=C(f,0,\varepsilon)$.

The following lemma is a direct analogue of [S3, Lemma 2] and so we give the proof only for completeness.

Lemma 2.1. *Let (X,ϱ) be a metric space whose square X^2 is a Baire space. Let $f\in C(X)$ and $\alpha\in\mathbb{R}$. Then the following three conditions are equivalent:*

(i) *$C_1(f,\alpha)$ is residual in X^2,*
(ii) *$C_1(f,\alpha)$ is dense in X^2,*
(iii) *for every two balls B_1, B_2, $\liminf_{n\to\infty}\varrho(f^n(B_1),f^n(B_2))\leqslant\alpha$.*

Proof. The implications $(i)\Rightarrow(ii)\Rightarrow(iii)$ are obvious. We are going to prove $(iii)\Rightarrow(i)$. So let (iii) be fulfilled. We have $C_1(f,\alpha)=\bigcap_{n=1}^{\infty}L(n,\alpha+\frac{1}{n})$ where

$$L\left(n,\alpha+\frac{1}{n}\right)=\left\{[x,y]\in X^2:\inf_{k\geqslant n}\varrho(f^k x,f^k y)<\alpha+\frac{1}{n}\right\}\ .$$

For every n, $L(n,\alpha+\frac{1}{n})$ is obviously an open set in X^2. To show that $C_1(f,\alpha)$ is residual it is thus sufficient to prove that for every n, $L(n,\alpha+\frac{1}{n})$ is dense in X^2. So fix n and balls B_1, B_2. We prove that $L(n,\alpha+\frac{1}{n})\cap(B_1\times B_2)\neq\emptyset$. From (iii) it follows that there exists $k\geq n$ with $\varrho(f^k(B_1),f^k(B_2))<\alpha+\frac{1}{n}$. This implies the existence of points $x\in B_1$, $y\in B_2$ such that $\varrho(f^k x,f^k y)<\alpha+\frac{1}{n}$. Hence $[x,y]\in L(n,\alpha+\frac{1}{n})$ and the proof is complete. $\square$

Next lemma shows that in case of the set $C_2(f,\beta)$ the situation in metric spaces is more complicated than the one on the interval and one can get only a weaker result than that from [S1, Lemma 4.16].

Lemma 2.2. *Let (X, ϱ) be a metric space whose square X^2 is a Baire space. Let $f \in C(X)$ and $\beta \in \mathbb{R}$. Consider the following conditions:*

(i) *$C_2(f, \beta)$ is residual,*
(ii) *$C_2(f, \beta)$ is dense,*
(iii) *for every ball B, $\limsup_{n\to\infty} \operatorname{diam} f^n(B) > \beta$,*
(iv) *$C_2(f, \beta^*)$ is residual for every $\beta^* < \frac{\beta}{2}$.*

Then (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv).

Proof. The implications $(i) \Rightarrow (ii) \Rightarrow (iii)$ are obvious. We are going to prove $(iii) \Rightarrow (iv)$. So let (iii) be fulfilled. Fix $\beta^* < \frac{\beta}{2}$. Put

$$C_2\left(f, n, \frac{\beta}{2}\right) = \left\{[x, y] \in X^2 : \sup_{k \geqslant n} \varrho(f^k x, f^k y) > \frac{\beta}{2}\right\} .$$

Then $\bigcap_{n=1}^{\infty} C_2\left(f, n, \frac{\beta}{2}\right) \subset C_2(f, \beta^*)$. For any n, $C_2\left(f, n, \frac{\beta}{2}\right)$ is open. Therefore to get (iv), it is sufficient to prove that for any n, the set $C_2\left(f, n, \frac{\beta}{2}\right)$ is dense in X^2. To this end, fix n and balls B_1, B_2. We need to show that $C_2\left(f, n, \frac{\beta}{2}\right) \cap (B_1 \times B_2) \neq \emptyset$.

Distinguish two cases.

Case 1. For some r, $f^r(B_1) \subset f^r(B_2)$. Since $\limsup_{i\to\infty} \operatorname{diam} f^i(B_1) > \beta$ we can take $k \geqslant \max\{r, n\}$ with $\operatorname{diam} f^k(B_1) > \beta$. Since $f^k(B_1) \subset f^k(B_2)$ there are $x \in B_1$, $y \in B_2$ with $\varrho(f^k x, f^k y) > \beta$ whence $[x, y] \in C_2(f, n, \beta) \subset C_2(f, n, \frac{\beta}{2})$.

Case 2. For every r, $f^r(B_1) \setminus f^r(B_2) \neq \emptyset$. Now take $k \geqslant n$ with $\operatorname{diam} f^k(B_2) > \beta$ and a point $u \in f^k(B_1) \setminus f^k(B_2)$. Then there is a point $v \in f^k(B_2)$ such that $\varrho(u, v) > \frac{\beta}{2}$, since otherwise for any two points $v_1, v_2 \in f^k(B_2)$ we would have $\varrho(v_1, v_2) \leqslant \varrho(v_1, u) + \varrho(u, v_2) \leqslant \beta$ and hence $\operatorname{diam}(B_2) \leqslant \beta$, a contradiction. Now take f^k-preimages $x \in B_1$ and $y \in B_2$ of u and v, respectively. Then $\varrho(f^k x, f^k y) > \frac{\beta}{2}$ and again $[x, y] \in C_2\left(f, n, \frac{\beta}{2}\right)$. $\square$

From Lemma 2.1 and Lemma 2.2 we get

Theorem 2.3. *Let (X, ϱ) be a metric space whose square X^2 is a Baire space. Let $f \in C(X)$ and α, $\beta \in \mathbb{R}$. Consider the following four conditions:*

(a) *f is generically (α, β)-chaotic,*
(b) *f is densely (α, β)-chaotic,*
(c) *the following two conditions are fulfilled simultaneously:*
  (c1) *for every two balls B_1, B_2, $\liminf_{n\to\infty} \varrho(f^n(B_1), f^n(B_2)) \leqslant \alpha$,*
  (c2) *for every ball B, $\limsup_{n\to\infty} \operatorname{diam} f^n(B) > \beta$,*
(d) *f is generically (α, β^*)-chaotic for every $\beta^* < \frac{\beta}{2}$.*

Then (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (d).

Now we are ready to prove Theorem A.

Proof of Theorem A. By putting $\alpha = 0$ and $\beta = \varepsilon$ in Theorem 2.3, we get Theorem A except for the claim that, in general, the condition (c) from Theorem A does not imply the generic ε-chaoticity of the map f.

To prove this claim fix $\varepsilon > 0$ and $a > 0$ and consider the following seven points in the euclidean plane: $V = [0,0]$, $A = [\frac{\varepsilon}{2} + a, 0]$, $B = [-\frac{\varepsilon}{2} - a, 0]$, $C = [0, -a]$, $C_1 = [0, -\frac{a}{4}]$, $C_2 = [0, -\frac{a}{2}]$, $C_3 = [0, -\frac{3}{4}a]$. Let X be the subspace of the euclidean plane defined as the union of the straight line segments AB and VC (i.e., X has the form of the letter T). Define a map $f : X \to X$ as follows. Let $f(V) = f(C_2) = f(C) = V$, $f(C_1) = A$, $f(C_3) = B$, $f(A) = f(B) = C$ and let f be affine on each of the straight line segments VC_1, C_1C_2, C_2C_3, C_3C, VA, VB. Then f is continuous, $f(AB) = VC$, $f(VC) = AB$ and for every ball G in X there is some n with $f^n(G) \supset AB$. Then $f^{n+1}(G) \supset VC$, $f^{n+2}(G) \supset AB$, etc. Hence the condition (c) from Theorem A is fulfilled.

On the other hand, repeat that $f(VC) = AB$ and $f(AB) = VC$ and notice that

$$L := \max\{\varrho(x,y) : x \in VC,\ y \in AB\} = \sqrt{a^2 + \left(\frac{\varepsilon}{2} + a\right)^2}\ .$$

Thus $\limsup_{n\to\infty} \varrho(f^n x, f^n y) \leqslant L$ whenever $x \in VC$ and $y \in AB$. For sufficiently small a we get $L \leqslant \varepsilon$ and in such a case $(VC \times AB) \cap C_2(f, \varepsilon) = \emptyset$. Consequently, f is not generically ε-chaotic. $\square$

Remark 2.4. Since in the proof of Theorem A we have $\lim_{a\to 0} L = \frac{\varepsilon}{2}$, the constant $\frac{\varepsilon}{2}$ at the very end of Theorem A cannot be replaced by any larger number — such a 'universal' (i.e., depending only on ε and not on the space under consideration) 'constant' larger than $\frac{\varepsilon}{2}$ does not exist. Nevertheless, for a *particular* space X it can happen that the constant $\frac{\varepsilon}{2}$ *can* be replaced by a larger number (and, even, a question is whether there is a space where this does not happen). For instance, in case of our 'letter T' space with fixed a we can replace $\frac{\varepsilon}{2}$ by any number smaller than $\frac{\varepsilon}{2} + a$. Moreover, using the idea from the proof of [S1, Lemma 4.15] one can even prove that this is the case when X is any finite graph. Still, our 'letter T' spaces show that there is no number larger than $\frac{\varepsilon}{2}$ which could serve as the mentioned 'universal' (depending only on ε) 'constant' for the class of all finite graphs.

3. Proof of Theorem B.

By ΔABC we will denote the triangle with vertices A, B, C (here we think of a triangle as a convex subset of the plane).

Recall that a map $f \in C(X)$ is called *exact* if for any ball B in X there exists $n \in \mathbb{N}$ with $f^n(B) = X$. An example of such a map is the standard tent map $\tau(x) = 1 - |2x - 1|$ defined on the unit interval $I = [0,1]$.

A continuous map $F \in C(I^2)$ is called triangular if it is of the form $F(x,y) = (f(x), g(x,y))$. Instead of $g(x,y)$ we also write $g_x(y)$. Here $\{g_x,\ x \in I\}$ is a family of continuous maps from $C(I)$ depending continuously on $x \in I$.

The following lemmas are intuitively obvious but for completeness we give proofs.

Lemma 3.1. *There is a triangular map $F(x,y) = (f(x), g_x(y))$ in $C(I^2)$ such that F is exact, g_0 and g_1 are the identity maps $I \to I$ and the set $I \times \{0\}$ is F-invariant.*

Proof. Put $f = \tau$, $g_0 = g_1 = \mathrm{id}$. For every $x \in [\frac{1}{4}, \frac{1}{2}]$ let g_x be the map such that $g_x(0) = g_x(\frac{2}{3}) = 0$, $g_x(\frac{1}{3}) = g_x(1) = 1$ and g_x is linear on each of the intervals $[0, \frac{1}{3}]$, $[\frac{1}{3}, \frac{2}{3}]$ and $[\frac{2}{3}, 1]$. Further, for $x \in [0, \frac{1}{4}]$ let g_x be the map uniquely determined by the following conditions:

- $g_x(0) = 0$, $g_x(\frac{1}{2}) = \frac{1}{2}$, $g_x(1) = 1$,

- g_x is piecewise linear with three pieces of linearity,
- the slope of g_x in a right neighbourhood of 0 as well as that in a left neighbourhood of 1 is $8x+1$ and the slope of g_x in a neighbourhood of $\frac{1}{2}$ is -3.

Finally, for every $x \in (\frac{1}{2}, 1]$ put $g_x = g_{1-x}$.

Obviously, F is well defined, continuous and the set $I \times \{0\}$ is F-invariant. Notice that for any x and any interval $J \subset I$, $\operatorname{diam} g_x(J) \geqslant \frac{1}{3} \operatorname{diam} J$.

We are going to prove that F is exact. So, take nondegenerate intervals $J_1, J_2 \subset I$. We need to show that there exists N with $F^N(J_1 \times J_2) = I^2$.

First take n with $\tau^n(J_1) = I$ and denote $S_x = \{y \in I : [x, y] \in F^n(J_1 \times J_2)\}$. If we denote $\delta = (\frac{1}{3})^n \operatorname{diam} J_2$ then one can see that for every x and every component s_x of S_x we have $\operatorname{diam} s_x \geqslant \delta > 0$.

Now take k such that for every interval $J \subset I$ whose length is at least δ, $(g_{\frac{2}{3}})^k(J) = I$. This together with the facts that the point $\frac{2}{3}$ is fixed for τ and for all x sufficiently close to $\frac{2}{3}$ we have $g_x = g_{\frac{2}{3}}$, imply that $F^k(F^n(J_1 \times J_2)) \supset [\frac{2}{3} - \varepsilon, \frac{2}{3} + \varepsilon] \times I$ for some $\varepsilon > 0$.

Finally, take r with $\tau^r([\frac{2}{3} - \varepsilon, \frac{2}{3} + \varepsilon]) = I$. Since all the maps g_x, $x \in I$ are onto, it is sufficient to put $N = n + k + r$. □

Lemma 3.2. *Given a triangle $T = \Delta ABC$, there is an exact map $f \in C(T)$ such that all the points from $AB \cup AC$ are fixed points of f.*

Proof. By Lemma 3.1 there is an exact triangular map $F(\varphi, r) = (\tau(\varphi), g_\varphi(r))$, $\varphi \in I$, $r \in I$ such that $g_0 = g_1 = \mathrm{id}$. Since the set $I \times \{0\}$ is F-invariant, we can think of φ and r as of polar coordinates. In such a way F becomes a continuous map from a disc sector $\{[\varphi, r] : \varphi \in [0,1], r \in [0,1]\}$ into itself. Obviously, F is exact and all the points of the form $[0, r]$ and $[1, r]$, $r \in [0,1]$ are fixed points of F. Using the topological conjugacy via an appropriate homeomorphism from the disc sector onto T we get a map $f \in C(T)$ with all the required properties. □

Now we are ready to prove Theorem B.

Proof of Theorem B. In the plane take the points given in polar coordinates φ, r by $V = [0, 0]$ and $A_n = \left[\frac{\pi}{2^n}, \frac{1}{2^{n-1}}\right]$, $n = 1, 2, \ldots$. Consider the set $X = \bigcup_{n=1}^{\infty} VA_n$, i.e. a union of straight line segments, endowed with the metric inherited from the euclidean plane. Obviously, X is a continuum.

Define $f \in C(X)$ as follows. Let $f(V) = V$ and for any n, let $f|_{VA_n}$ be topologically conjugate to the tent map.

Since $f(VA_n) = VA_n$ and the set $VA_n \setminus \{V\}$ is open in X, the fact that $\operatorname{diam}(VA_n) \to 0$ when $n \to \infty$ shows that the condition (c2) from Theorem A is not fulfilled for any $\varepsilon > 0$. Hence f is not generically ε-chaotic for any $\varepsilon > 0$.

We are going to show that f is generically chaotic. To this end denote $X_k = \bigcup_{n=1}^{k} VA_n$, $k \in \mathbb{N}$ and realize that the exactness of the tent map gives the exactness of $f|_{VA_n}$ for every n. This implies that for any ball B in X, $f^r(B) \supset VA_s$ for some r and s. Hence, by Theorem A, for any fixed k the map $f|_{X_k}$ is generically ε_k-chaotic for some $\varepsilon_k > 0$. Therefore the set M_k of points from X_k^2 which are not Li-Yorke pairs, is of first category in X_k^2 and hence of first category in X^2. Since any point from X^2 belongs to X_k^2 for some k, we then get that the set of points from X^2 which are not Li-Yorke pairs is the first category set $\bigcup_{k=1}^{\infty} M_k$. Thus f is generically chaotic.

Now we are going to modify the described example in order that the space be convex.

Denote $T_n = \Delta VA_nA_{n+1}$, $n = 1, 2, \dots$. Then $Y = \bigcup_{n=1}^{\infty} T_n$ is a convex continuum in the plane. By Lemma 3.2, for every n there is an exact map $g_n \in C(T_n)$ such that every point from $VA_n \cup VA_{n+1}$ is a fixed point of g_n. Let g be a self-map of Y defined as follows. For $y \in Y$ put $g(y) = g_k(y)$ where k is such that $y \in T_k$. It is easy to see that g is well defined and continuous. To prove that g is generically chaotic but not generically ε-chaotic for any $\varepsilon > 0$, repeat the above proof that the map f has these properties (just replace VA_n by T_n and X_k by $Y_k = \bigcup_{n=1}^{k} T_n$). □

Added in proof. After submitting the paper the author learned about the recent preprint [HY] which is written in the setting of compact metric spaces and surjective maps and which partially overlaps with the present paper (cf. our Theorem A and the equivalence of (2), (3) and (4) in the 'sensitive' case of Theorem 3.5 from [HY]).

References

[HMcC] R. C. Haworth and R. A. McCoy, *Baire spaces*, Dissertationes Mathematicae **141** (1977), 1 – 77.

[HY] W. Huang and X. Ye, *Devaney's chaos or 2-scattering implies Li-Yorke's chaos*, Preprint.

[KS] S. Kolyada and Ľ. Snoha, *Some aspects of topological transitivity – a survey*, Grazer Math. Ber., Bericht **Nr. 334** (1997), 3 – 35.

[Kr] M. R. Krom, *Cartesian products of metric Baire spaces*, Proc. Amer. Math. Soc. **42** (1973), 588 – 594.

[LY] T.-Y. Li and J. A. Yorke, *Period three implies chaos*, Amer. Math. Monthly **82** (1975), 985 – 992.

[P] J. Piórek, *On the generic chaos in dynamical systems*, Acta Math. Univ. Iagell. **25** (1985), 293 – 298.

[S1] Ľ. Snoha, *Generic chaos*, Comment. Math. Univ. Carolinae **31** (1990), 793 – 810.

[S2] Ľ. Snoha, *Dense chaos*, Comment. Math. Univ. Carolinae **33** (1992), 747 – 752.

[S3] Ľ. Snoha, *Two–parameter chaos*, Acta Univ. M. Belii, Ser. Math. **No.1** (1993), 3 – 6.

(Received May 13, 2000)

Dept. of Mathematics
Faculty of Finance
Matej Bel University
Tajovského 10
974 01 Banská Bystrica
SLOVAKIA

E-mail address: murinova@financ.umb.sk

Acta Univ. M. Belii
Math. no. 8(2000), pp. 51-57

ON RELATIONS SATISFYING SOME HORN FORMULAS

Branimir Šešelja and Andreja Tepavčević

Abstract. A general approach to relations usually considered on a set is presented. Relations are supposed to satisfy particular Horn formulas. It is proved that this approach is equivalent to the particular framework developed for investigation of compatible relations on algebras. Collections of such relations are algebraic lattices under inclusion, with an ideal being isomorphic with the power set of A. Conditions are presented under which such a lattice consists of all relations which are reflexive on subsets of A. These conditions turn out to be closely connected with lattice properties of the diagonal relation on A.

1 Introduction

It is known that an algebra $\mathcal{A}$ determines particular lattices of $\mathcal{A}$-compatible binary relations, such as congruence and tolerance lattices, not only on $\mathcal{A}$, but also on each subalgebra of $\mathcal{A}$. In [2], a framework for the generation of such lattices was introduced.

In the present paper, another approach to these algebraic lattices is developed, and some new results are proved. Starting with a set A, we consider all binary relations on A which satisfy a set of particular Horn formulas. We prove that relations satisfying these Horn formulas on a set are precisely those which are introduced in [2] for algebras. Conditions which should be satisfied by the Horn formulas, in order that diagonal relations and also some other connected relations belong to the collection are given. Further, there is a Horn formula whose presence provides the existence of a congruence on the lattice of relations, such that its blocks consist of reflexive relations on subsets of A. We prove that properties of the diagonal relation yield some structural properties of the corresponding lattice.

2 Results

Let $\mathcal{L}$ be a first order language with only one relational symbol α which is binary, and with no functional symbols. Let $\mathcal{S}$ be a set of universal formulas of the type $\mathcal{L}$ over a set of variables X, such that each $\varphi \in \mathcal{S}$ is as follows:

$$\varphi \equiv (\forall x_1)\dots(\forall x_k)(F_1\&\dots\&F_n \Longrightarrow G_1\&\dots\&G_m), \tag{1}$$

2000 *Mathematics Subject Classification.* 04A05, 08A05.
Key words and phrases. Relation on a set, tolerance, quasi-order, equivalence relation, weak reflexivity, algebraic lattice.

where F_i, G_j, $i = 1, \dots, n, j = 1, \dots, m$ are atomic formulas (i.e., of the form $\alpha(x_p, x_q)$, $p, q \in \{1, \dots, k\}$) and the set of variables occurring in $G_1, \dots, G_m$ is a subset of the set of variables occurring in $F_1, \dots, F_n$.

If $\mathcal{S}$ is a foregoing set of formulas and A a nonempty set, then denote by $\mathcal{R}_{\mathcal{S}}^A$ the set of all relations ρ on A, such that $(A, \rho) \models \mathcal{S}$.

Let $\Sigma = \{(\sigma_i, \sigma_i') : i \in I\}$ where I is an index set and for each $i \in I$ both σ_i and σ_i' are relations on the same set V_i, satisfying:

$$(2) \quad \{x : x\sigma_i' y \text{ or } y\sigma_i' x, \text{ for some } y \in V_i\} \subseteq \{x : x\sigma_i y \text{ or } y\sigma_i x, \text{ for some } y \in V_i\}.$$

Denote by R_Σ^A the set of all relations ρ such that for all $i \in I$

$$(3) \qquad Hom(\sigma_i, \rho) \subseteq Hom(\sigma_i', \rho),$$

(where Hom (γ, δ) is the set of all relational homomorphisms from a relation γ on C into a relation δ on D; i.e., maps $f : C \to D$ such that from $a\gamma b$ it follows that $f(a)\delta f(b)$).

Theorem 1. *Let A be a set. If R_Σ^A is the collection of relations described as above, where $|\sigma_i| < \aleph_0$, then there is a set of Horn formulas $\mathcal{S}$, such that the set $\mathcal{R}_{\mathcal{S}}^A$ coincides with R_Σ^A. Conversely, if $\mathcal{S}$ is a set of Horn formulas described at the beginning, then there are sets V_i and Σ defined above, such that the collection $\mathcal{R}_{\mathcal{S}}^A$ coincides with R_Σ^A.*

Proof. From the condition $|\sigma_i| < \aleph_0$, it follows that $|\sigma_i'| < \aleph_0$ (by (2)). We can also assume that each V_i is finite. To every ordered pair (σ_i, σ_i') of relations on a set V_i, for $|V_i| < \aleph_0$ there corresponds a Horn formula, as described in the sequel, such that a relation ρ on the set A satisfies that formula if and only if it satisfies (3).

Let h be a bijection between V_i and a set of variables $X = \{x_1, x_2, \dots, x_k\}$. Further, for each pair $(a_j, b_j) \in \sigma_i$, $j \in J = \{1, \dots, n\}$ consider atomic formula $F_j = \alpha(h(a_j), h(b_j))$. Similarly, for each pair $(c_l, d_l) \in \sigma_i'$, $l \in K = \{1, \dots, m\}$, consider atomic formula $G_l = \alpha(h(c_l), h(d_l))$.

Now, let A be a set and $\rho \subseteq A^2$ a relation which satisfies condition (3). If $f : V_i \longrightarrow A$, and $(a_j, b_j) \in \sigma_i$ implies that $(f(a_j), f(b_j)) \in \rho$, then for the same f, from $(a_j, b_j) \in \sigma_i'$ it follows that $(f(a_j), f(b_j)) \in \rho$. The mapping $\mathcal{V} = h^{-1} \circ f$ maps X to A i.e., it is a valuation. Thus, every mapping f corresponds to a valuation. On the other hand, if $\mathcal{V} : X \longrightarrow A$ is a valuation, then $f = h \circ \mathcal{V}$ is the corresponding mapping from V_i to A. From the previous consideration it follows that every f from Hom(σ_i, ρ) also belongs to Hom(σ_i', ρ) if and only if the Horn formula that corresponds to (3) for $\alpha = \rho$, is true in every valuation.

Further, observe a Horn formula φ as in (1), where V is a set of variables appearing in it. Let σ and σ' be relations on V defined by:

$(x, y) \in \sigma$ if and only if there is an atomic formula $F_i = \alpha(x, y)$ in the antecedent of φ and

$(x, y) \in \sigma'$ if and only if there is an atomic formula $G_i = \alpha(x, y)$ in the consequent of φ.

By the consideration as above we conclude that the relations on A satisfying the formula φ and the corresponding inclusion (3) coincide. □

Corollary 1. *The set $\mathcal{R}_{\mathcal{S}}^{A}$ is an algebraic lattice under inclusion.*

Proof. Consider a trivial algebra $\mathcal{A} = (A, f)$ $(f(x) = x)$ on $A \neq \emptyset$. Then, $\mathcal{R}_{\mathcal{S}}^{A}$ contains compatible relations. Since the corresponding (according to Theorem 1) collection of relations R_{Σ}^{A} is an algebraic lattice under inclusion by Proposition 2 in [2] (because it coincides with R_{Σ}^{A}), $\mathcal{R}_{\mathcal{S}}^{A}$ is also an algebraic lattice. □

In the sequel, we consider a set of Horn formulas $\mathcal{S}$ such that the diagonal relation on a given set satisfies each of them.

Let $\mathcal{S}$ be a set of Horn formulas of the type (1), such that

$$(A, \Delta_A) \models \mathcal{S} \tag{4}$$

holds for a nonempty set A (Δ_A, or Δ is the diagonal relation on A).

Next we describe Horn formulas which satisfy (4).

Observe that every formula defined by (1) over a set of variables X is equivalent to the finite conjunction of formulas ϕ being of the form

$$\phi \equiv (\forall x_1, \ldots, x_k)(\alpha(x_{i_1}, x_{i_2})\& \ldots \&\alpha(x_{i_{p-1}}, x_{i_p}) \Longrightarrow \alpha(x_1, x_2)), \tag{5}$$

where $\{x_{i_1}, \ldots, x_{i_p}\} = X_\phi = \{x_1, \ldots, x_k\} \subseteq X$.

Denote by $\mathcal{T}_\phi$ the set of atomic formulas figuring in the antecedent of ϕ:

$\mathcal{T}_\phi = \{\alpha(x_{i_1}, x_{i_2}), \ldots, \alpha(x_{i_{p-1}}, x_{i_p})\}$.

Further on, for every $x \in X_\phi$, we define $U_\phi(x)$, as follows:

$y \in U_\phi(x)$ if and only if there are $n \in N$ and $u_0, \ldots, u_n \in X_\phi$, such that $\alpha(u_j, u_{j+1}) \in \mathcal{T}_\phi$ or $\alpha(u_{j+1}, u_j) \in \mathcal{T}_\phi$, for $j = 0, ..., n-1$ and $x = u_0$, $y = u_n$.

A part of the following proposition is a consequence of Lemma 2 in [2], but we provide another proof.

Proposition 1. *Let A be a set, such that $|A| > 1$ and ϕ a Horn formula defined by (5) over a set of variables X. Then, $\Delta_B \models \phi$ for all $\emptyset \neq B \subseteq A$ if and only if $U_\phi(x_1) \cap U_\phi(x_2) \neq \emptyset$.*

Proof. Suppose that $U_\phi(x_1) \cap U_\phi(x_2) = \emptyset$. Then, let B be a nonempty subset of A and $a, b \in B$, $a \neq b$. Consider the valuation $\mathcal{V} : X_\phi \mapsto A$, such that $\mathcal{V}(u) = a$ for every $u \in U_\phi(x_1)$, and $\mathcal{V}(v) = b$ for every $v \notin U_\phi(x_1)$. Then obviously Δ_B is not a model for ϕ, since all atomic formulas from $\mathcal{T}_\phi$ in this interpretation are associated to ordered pairs with equal coordinates ((a, a) or (b, b)) belonging to Δ_B, and $\alpha(x_1, x_2)$ is interpreted by $(a, b) \notin \Delta_B$.

Conversely, let $U_\phi(x_1) \cap U_\phi(x_2) \neq \emptyset$. Then the diagonal relation of any nonempty subset B of A is a model of ϕ. Indeed, for any valuation $\mathcal{V} : X_\phi \to A$ which assigns different values $a, b \in B$ to x_1 and x_2, both antecedent and consequent of ϕ are false if α is interpreted by Δ_B, hence ϕ is satisfied. Obviously, ϕ is satisfied also in the case when the same value from B is assigned to x_1 and x_2. Thus, $\Delta_B \models \phi$.

Observe that the empty set trivially satisfies any set of Horn formulas of this type, hence $\emptyset \in \mathcal{R}_{\mathcal{S}}^{A}$, for every nonempty set A. □

Proposition 2. *Let A, $\mathcal{S}$ and $\mathcal{R}^A_{\mathcal{S}}$ be as above. Then, the principal ideal $\Delta\downarrow$ in $\mathcal{R}^A_{\mathcal{S}}$ is isomorphic with the power set of A.*

Proof. By the condition (4), diagonal relations of all subsets of A belong to $\mathcal{R}^A_{\mathcal{S}}$. Hence, the mapping $f : \Delta_B \mapsto B$, $B \subseteq A$ (with $\Delta_\emptyset = \emptyset$) is obviously the required isomorphism. □

Proposition 3. *If B is a nonempty subset of A, then (i) B^2 and (ii) $B^2 \cup \Delta$ are relations from $\mathcal{R}^A_{\mathcal{S}}$*

Proof. Let $\phi \in \mathcal{S}$, as described by (5):
$\phi \equiv (\forall x_1, \ldots, x_k)(\alpha(x_{i_1}, x_{i_2})\& \ldots \& \alpha(x_{i_{p-1}}, x_{i_p}) \Longrightarrow \alpha(x_1, x_2))$.

(i) If the antecedent of ϕ is satisfied by B^2, then obviously the consequent also holds. Indeed, by the assumption the variables x_1 and x_2 appear also in the antecedent of ϕ. Therefore, B^2 is a model of ϕ.

(ii) Suppose that for any valuation, $B^2 \cup \Delta$ satisfies the antecedent of ϕ, i.e., that the interpretation of $\alpha(x_{i_m}, x_{i_{m+1}})$ is either $(a, a) \in \Delta$, $a \in A$, or $(b, c) \in B^2$, $b, c \in B$. If the interpretation of $\alpha(x_1, x_2)$ is an ordered pair (d, e) from B^2, then $B^2 \cup \Delta \models \phi$. If one of these coordinates, e.g. d, is not an element from B, then, since $\Delta \in \mathcal{R}^A_{\mathcal{S}}$, by Proposition 1 it follows that $d = e$. Thus again $B^2 \cup \Delta \models \phi$.

Hence, $B^2 \cup \Delta \in \mathcal{R}^A_{\mathcal{S}}$. □

In the sequel, Δ is supposed to belong to $\mathcal{R}^A_{\mathcal{S}}$, for every A. We discuss particular cases of such lattices, examples of which are well known.

If a relation $\rho \subseteq A^2$ satisfies the formula

$$\Phi \equiv (\forall x)(\forall y)(\alpha(x, y) \Rightarrow \alpha(x, x)\&\alpha(y, y)), \tag{6}$$

then it is called a **weakly reflexive** relation on A.

Some particular known cases are as follows. Let A be a nonempty set and $Rw\,A$ the set of all weakly reflexive relations on A; $Qw\,A$ the set of all relations on A which are reflexive and transitive on subsets of A (weak quasi-orders on A); $Tw\,A$ the set of all relations on A which are reflexive and symmetric on subsets of A (weak tolerances on A); $Ew\,A$ the set of all relations on A which are symmetric and transitive on subsets of A (weak equivalences on A). Obviously, all these relations satisfy the formula (6).

It is easy to see that all the mentioned sets are algebraic lattices of the form $\mathcal{R}^A_{\mathcal{S}}$, for a suitable set of Horn formulas $\mathcal{S}$. Hence, in all these lattices the principal ideal $\Delta\downarrow$ generated by the diagonal relation Δ on A is isomorphic with the power set $\mathcal{P}(A)$ of A. However, these lattices have some additional properties, as follows. The filter $\Delta\uparrow$ (i.e., the interval-sublattice $[\Delta, A^2]$) is the lattice of the corresponding reflexive relations on the whole set A. Each of these is a disjoint union of interval lattices $[\Delta_B, B^2]$, $B \subseteq A$.

Next we give conditions under which $\mathcal{R}^A_{\mathcal{S}}$ has the foregoing properties.

Recall that $a \in L$ is said to be **codistributive** if for all $x, y \in L$, $a \wedge (x \vee y) = (a \wedge x) \vee (a \wedge y)$. Such element induces a homomorphism n_a of L onto $a\downarrow$, defined by $n_a(x) = x \wedge a$.

Observe that in an algebraic lattice every codistributive element is infinitely codistributive ([5]). In this case, the congruence classes induced by n_a have maximal elements.

In all algebraic lattices listed above, the diagonal relation Δ is an (infinitely) codistributive element. Moreover, maximal elements of the congruence classes induced by n_Δ are squares of subsets of $\mathcal{A}$.

Theorem 2. *Let $\mathcal{S}, A$ and $\mathcal{R}_{\mathcal{S}}^A$ be as above. The following are equivalent:*

(i) Δ is a codistributive element in $(\mathcal{R}_{\mathcal{S}}^A, \subseteq)$, and maximal elements of the congruence classes induced by n_Δ are squares of subsets of A;

(ii) $\mathcal{S} \vdash \Phi$, where Φ is given by (6);

(iii) $\mathcal{R}_{\mathcal{S}}^A$ is a disjoint union of lattices consisting of reflexive relations on subsets of A, which satisfy $\mathcal{S}$.

Proof. $(ii) \Rightarrow (i)$ follows by Proposition 3 in [2].

$(i) \Rightarrow (ii)$ Suppose that there is a relation $\rho \in \mathcal{R}_{\mathcal{S}}^A$ which does not satisfy (ii), i.e., such that for some $a, b \in A$
$(a,b) \in \rho$ and at most one of two pairs (a,a), (b,b) is in ρ.

Take $(a,b) \in \rho$ and $(a,a) \notin \rho$, and let $B = \{x \in A \mid (x,x) \in \rho\}$.

Now, if Δ is a codistributive element in $\mathcal{R}_{\mathcal{S}}^A$, then, since $\rho \wedge \Delta = \Delta_B$, it follows that ρ belongs to the same class of the congruence induced by n_Δ as B^2. However, $\rho \not\leq B^2$ and B^2 is not the greatest element of the class.

$(ii) \Rightarrow (iii)$ Suppose that every $\rho \in \mathcal{R}_{\mathcal{S}}^A$ satisfies the formula Φ, i.e., that

$$(x,y) \in \rho \text{ implies } (x,x) \in \rho \text{ and } (y,y) \in \rho.$$

Then, for $B = \{x \mid (x,x) \in \rho\}$, $\Delta_B \in \mathcal{R}_{\mathcal{S}}^A$, and $B^2 \in \mathcal{R}_{\mathcal{S}}^A$. In addition,

$$\Delta_B \leq \rho \leq B^2 \tag{7}$$

i.e., $\mathcal{R}_{\mathcal{S}}^A = \bigcup([\Delta_B, B^2] \mid B \subseteq A)$.

$(iii) \Rightarrow (i)$ If $\rho, \theta \in \mathcal{R}_{\mathcal{S}}^A$, then there are $B, C \subseteq A$, such that $\rho \in [\Delta_B, B^2]$, $\theta \in [\Delta_C, C^2]$. Now,

$$\rho \vee \theta \in [\Delta_{B \cup C}, (B \cup C)^2]$$

(since $\Delta_B \vee \Delta_C \leq \rho \vee \theta \leq B^2 \vee C^2 \leq (B \cup C)^2$).

Hence, $(\rho \vee \theta) \wedge \Delta = \Delta_{B \vee C} = \Delta_B \vee \Delta_C = (\rho \wedge \Delta) \vee (\theta \wedge \Delta)$, which proves that Δ is a codistributive element of $(\mathcal{R}_{\mathcal{S}}^A, \subseteq)$.

By (7), B^2 is the greatest element of the class to which ρ belongs, since $\rho \wedge \Delta = B^2 \wedge \Delta = \Delta_B$. □

From now on, we assume that formula Φ given by (6) (describing the weak reflexivity) is a consequence of formulas in $\mathcal{S}$.

Proposition 4. *If $\rho \in \mathcal{R}_{\mathcal{S}}^A$, then also $\rho \cup \Delta \in \mathcal{R}_{\mathcal{S}}^A$.*

Proof. Let $\rho \in \mathcal{R}_{\mathcal{S}}^A$ and $\rho \cap \Delta = \Delta_B$, $B \subseteq A$. By Theorem 2, $\rho \leq B^2$. We have to prove that $\rho \vee \Delta = \rho \cup \Delta$ in $\mathcal{R}_{\mathcal{S}}^A$. Observe that $\rho \vee \Delta \leq B^2 \vee \Delta = B^2 \cup \Delta$, by Proposition 3. Now, if $\rho \cup \Delta \notin \mathcal{R}_{\mathcal{S}}^A$, then there is a formula (5), which is not satisfied by $\rho \cup \Delta$, i.e., there is a valuation $\mathcal{V}$ such that the antecedent is true, while the consequent is false. Let $\mathcal{V}(x_1) = b$, and $\mathcal{V}(x_2) = c$, in this valuation. Now, $b, c \in B$, $b \neq c$ and $(b,c) \notin \rho$. Since $\Delta \in \mathcal{R}_{\mathcal{S}}^A$, by Proposition 1, we have that $U_\phi(x_1) \cap U_\phi(x_2) \neq \emptyset$. In other words, in this valuation all elements from $U_\phi(x_1)$ and $U_\phi(x_2)$ have values from B. Now, starting with $\mathcal{V}$, we consider another valuation $\mathcal{V}'$ on B, as follows. Values of all variables which in $\mathcal{V}$ are elements from B remain the same while all other variables take the same value (from B). In this valuation ρ does not satisfy the formula, which gives a contradiction.

Hence, $\rho \cup \Delta \in \mathcal{R}_{\mathcal{S}}^{A}$, whenever $\rho \in \mathcal{R}_{\mathcal{S}}^{A}$. □

Next we prove that some properties of Δ enable structural decomposition of the lattice $\mathcal{R}_{\mathcal{S}}^{A}$.

As it is known, an element a of a bounded lattice L is **neutral** if the mappings $x \mapsto x \wedge a$ and $x \mapsto x \vee a$ are homomorphisms on L, and $x \mapsto (x \wedge a, x \vee a)$ is an embedding from L into $a\downarrow \times a\uparrow$.

Theorem 3. *A lattice identity holds on the lattice $\mathcal{R}_{\mathcal{S}}^{A}$ if and only if it holds on its sublattice $\Delta\uparrow$ of all reflexive relations from $\mathcal{R}_{\mathcal{S}}^{A}$.*

Proof. In every lattice $\mathcal{R}_{\mathcal{S}}^{A}$, Δ is a neutral element. This is an easy consequence of Proposition 4. The proof of the Theorem is now straightforward, by the definition of a neutral element, and by the fact that $\Delta\downarrow = \mathcal{P}(A)$. □

If $\mathcal{A} = (A, F)$ is an algebra, then $\mathcal{R}_{\mathcal{S}}^{\mathcal{A}}$ is the set of all relations from $\mathcal{R}_{\mathcal{S}}^{A}$ which are compatible with all fundamental operations on $\mathcal{A}$.

The following are almost immediate consequences of the above results.

Corollary 2. *Let $\mathcal{A} = (A, F)$ be an algebra and $\mathcal{S}$ a set of formulas as previously defined. Let also $\mathcal{R}_{\mathcal{S}}^{\mathcal{A}}$ be the set of all compatible relations on $\mathcal{A}$ which satisfy $\mathcal{S}$. Then $\mathcal{R}_{\mathcal{S}}^{\mathcal{A}}$ is an algebraic lattice under inclusion whose ideal $\Delta\downarrow$ is isomorphic with the lattice $Sub\mathcal{A}$.* □

Corollary 3. *If $\mathcal{A}$, $\mathcal{S}$ and $\mathcal{R}_{\mathcal{S}}^{\mathcal{A}}$ are as in Corollary 2, then the following are equivalent:*

(i) Δ is a codistributive element of the lattice $\mathcal{R}_{\mathcal{S}}^{\mathcal{A}}$ and the maximal elements of congruence classes induced by n_Δ are squares of subalgebras;

(ii) every $\rho \in \mathcal{R}_{\mathcal{S}}^{\mathcal{A}}$ is weakly reflexive;

(iii) $\mathcal{R}_{\mathcal{S}}^{\mathcal{A}}$ is a disjoint union of lattices consisting of reflexive, compatible relations on subalgebras of $\mathcal{A}$, which satisfy $\mathcal{S}$. □

References

[1] I. Chajda, *Algebraic theory of tolerance relations* (1991), Univerzita Palackého Olomouc.

[2] I.Chajda, B.Šešelja, A.Tepavčević, *Lattices of compatible relations satisfying a set of formulas*, Algebra Univers. 40 (1998), 51-58.

[3] G. Grätzer, *General lattice theory*, Akademie-Verlag Berlin, 1978.

[4] A.G. Pinus, I. Chajda, *O kvazipojadkah na universaljnih algebrah*, Algebra i logika **32 N3** (1993), 308-325.

[5] A. Tepavčević, *Diagonal relation as a continuous element in a weak congruence lattice*, Proc. of the Conf. "General Algebra and Ordered Sets", Olomouc (1994), 156-163.

[6] G. Vojvodić, B. Šešelja, *On the Lattice of Weak Congruence Relations*, Algebra Universalis **25** (1989), 221-230.

(Received January 26, 2000)

Institute of Mathematics
Fac. of Sci., University of Novi Sad
Trg D. Obradovića 4
21000 Novi Sad
YUGOSLAVIA

E-mail address: etepavce@eunet.yu